

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА

по направлению «Экономика»

ЭКОНОМЕТРИЧЕСКИЕ МЕТОДЫ ИЗУЧЕНИЯ МАКРОЭКОНОМИЧЕСКИХ
МОДЕЛЕЙ

Выполнил:

Обучающийся 4 курса, группы МиСМЭ-4,
Валова Мария Игоревна

_____/Подпись/

Научный руководитель:

Кандидат экономических наук, доцент
Евстратчик Светлана Васильевна

_____/Подпись/

Содержание

Введение	3
Глава I Экономика России	5
1.1 Региональное неравенство округов РФ	5
1.2 Причины и возможные последствия регионального неравенства.....	10
Выводы по главе I.....	12
Глава II Регрессионный и панельный анализ.....	13
2.1 Системы одновременных уравнений.....	13
2.1.1 Методы оценивания	16
2.1.2 Инструментальные переменные. Тест Хаусмана.....	22
2.2 Построение и оценивание СОУ	24
2.3 Модели панельных данных	29
2.3.1 Базовая модель панельных данных	29
2.3.2 Модели с фиксированными и случайными эффектами	31
2.3.3 Выбор наилучшей модели	36
2.4 Построение и анализ моделей панельных данных.....	38
2.5 Выбор наиболее предпочтительного округа.....	41
Выводы по главе II	49
Заключение	50
Список использованной литературы	52
Приложение 1 Система уравнений 2	53
Приложение 2 Система уравнений 3	54
Приложение 3 Система уравнений 4	55
Приложение 4 Система уравнений 5	56
Приложение 5 Система уравнений 6	57
Приложение 6 Система уравнений 7	58
Приложение 7 Система уравнений 8	59
Приложение 8 Модель с фиксированными эффектами.....	60
Приложение 9 Модель со случайными эффектами.....	61
Приложение 10 Данные	62
Приложение 11 Данные	65

Введение

Российская Федерация является крупнейшим по площади мировым государством, соответственно, не может не быть различий в уровне развития отдельных регионов. Данная особенность присуща большинству стран, имеющих крупные географические территории. По данным исследователей из Кембриджа¹, В 2013 году Россия занимала третье место в мировом рейтинге по неравенству регионов.

Целью выпускной работы является попытка выявления возможного влияния географического положения на развитие регионов Российской Федерации, выявление факторов, влияющих на неоднородное развитие округов РФ с помощью эконометрических инструментов.

В соответствии с поставленной целью требуется выполнить ряд задач:

1. Изучить вопрос неравномерности развития регионов РФ;
2. Анализировать литературу в области систем одновременных уравнений и панельных данных;
3. Собрать и систематизировать статистические данные;
4. Построить модели и провести анализ.

Объектом исследования являются федеральные округа Российской Федерации, а предметом – социально-экономическое неравенство федеральных округов РФ.

Вначале будут рассмотрены экономические аспекты выбранной темы: экономика России в 2000-2014 годах, анализ регионов, причины и возможные последствия неравенства регионов. Затем будут рассмотрены способы и методы возможного выявления закономерностей.

Описание сложных экономических явлений и процессов невозможно выполнить с помощью только одного эконометрического уравнения, зачастую бывает не достаточно построить даже множественную модель регрессии, ввиду того, что она не даёт значимых результатов, или же эти результаты абсурдны. К примеру, такие модели, как макроэкономическая модель Кейнса, Вальраса, Леонтьева описываются с помощью нескольких уравнений, т.е. систем. Следовательно, чтобы изучать и оценивать такие сложные модели, требуется владеть эконометрическим анализом систем одновременных уравнений. Таким образом, в работе будут рассмотрены такие инструменты анализа экономической ситуации в России, как основы регрессионного анализа, в частности,

¹ <http://www.finmarket.ru/main/article/3303366>

систем одновременных уравнений и основы панельных данных, т.к. зачастую, для более полного анализа необходимо оценивать данные по множеству объектов за некоторый период времени. В качестве практической реализации, будет построена и проанализирована система одновременных уравнений, а также будут изучены модели анализа панельных данных для Федеральных округов России. На основе построенных моделей будут сформулированы основные выводы и заключения.

Данные по федеральным округам, которые использовались для анализа, были получены с помощью Федеральной службы государственной статистики. В работе были использованы средства Excel, а также прикладной программный пакет Gretl, в котором производились основные расчёты и построение моделей.

Глава I Экономика России

1.1 Региональное неравенство округов РФ

Наша страна является крупнейшим по площади мировым государством, соответственно, необходимы административно-территориальные образования, которые разделяют территорию всего государства на ряд зон. Данное разделение необходимо для проведения успешной внутренней политики государства, проведения контроля, безопасности, а также для улучшения экономического, социального и политического положения каждого субъекта.

Федеральные округа РФ были организованы по Указу президента Российской Федерации В. В. Путина 13 мая 2000 года. Ф.О. были созданы по аналогии с экономическими районами и военными округами. На данный момент в РФ существует девять Федеральных Округов (в т ч Крымский). В работе не будет рассмотрен Крымский ФО, ввиду того, что данный округ не входил в состав Российской Федерации до 2014 года.



Рисунок 1.1 Федеральные округа Российской Федерации²

В таблице 1.1 представлены основные характеристики рассматриваемых Федеральных округов – суммарная площадь, количество входящих субъектов РФ, а также – административный центр.

² Федеральные округа РФ: <http://see2me.ru/karta-federalnyh-okrugov-rossii-s-krymom.html>

Таблица 1.1. Федеральные округа России³

№	Название Ф.О.	Площадь (км ²)	Субъектов РФ	Административный центр
1	Центральный	652 800	18	Москва
2	Южный	416 840	6	Ростов-на-Дону
3	Северо-Западный	1 677 900	11	Санкт-Петербург
4	Дальневосточный	6 215 900	9	Хабаровск
5	Сибирский	5 114 800	12	Новосибирск
6	Уральский	1 788 900	6	Екатеринбург
7	Приволжский	1 038 000	14	Нижний Новгород
8	Северо-Кавказский	172 360	7	Пятигорск

Одной из значимых характеристик Федеральных округов является показатель Валового Регионального Продукта.

Валовой Региональный Продукт [млн.руб] – это стоимость произведенных услуг и товаров резидентами данного региона для конечного пользования. Расчёт производится как в основных, так и в постоянных ценах. Данный показатель считается важным не только с точки зрения сравнения потребления и производства ВРП, но и со стороны рассмотрения уровня благосостояния жителей данного субъекта.

Анализируя показатели ВРП за период для различных округов, представленные на рисунке 1.2 можно заметить существенную разницу в абсолютных значениях. Такое различие в уровнях ВРП в разных регионах достигает до 19 триллионов рублей.

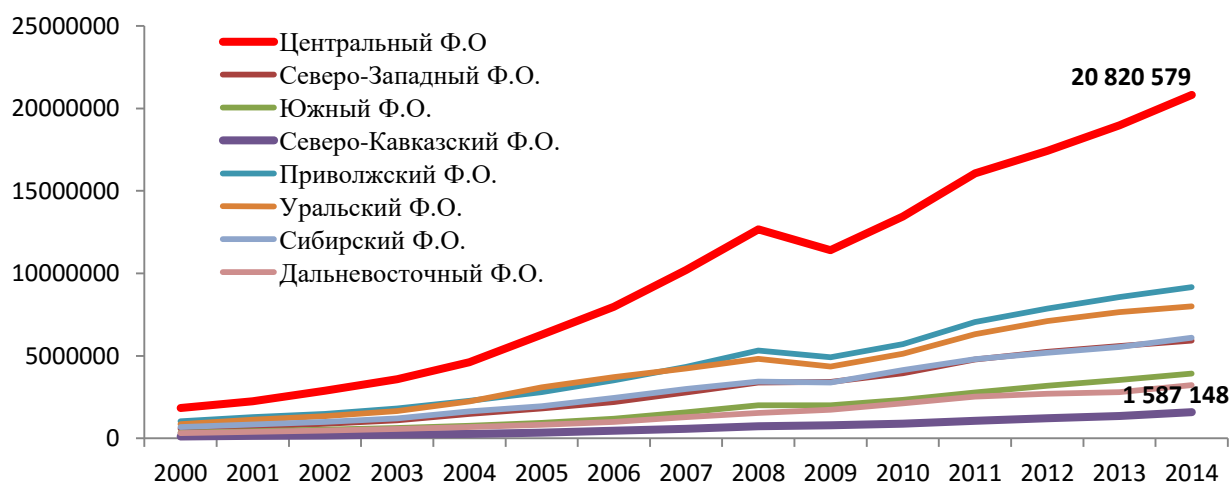


Рисунок 1.2 ВРП в Федеральных округах

Можно заметить, что тенденции развития – общие, т.е. в целом, тренды развития одинаковы.

³ Виртуальная Россия: <http://virru.ru/federalnye-okruga-rossii>

На рисунке 1.3 представлен суммарный ВВП по всем регионам. Видно, что лишь в 2008-2009 году отмечается снижение ВВП, что можно объяснить Мировым Финансовым Кризисом тех лет.

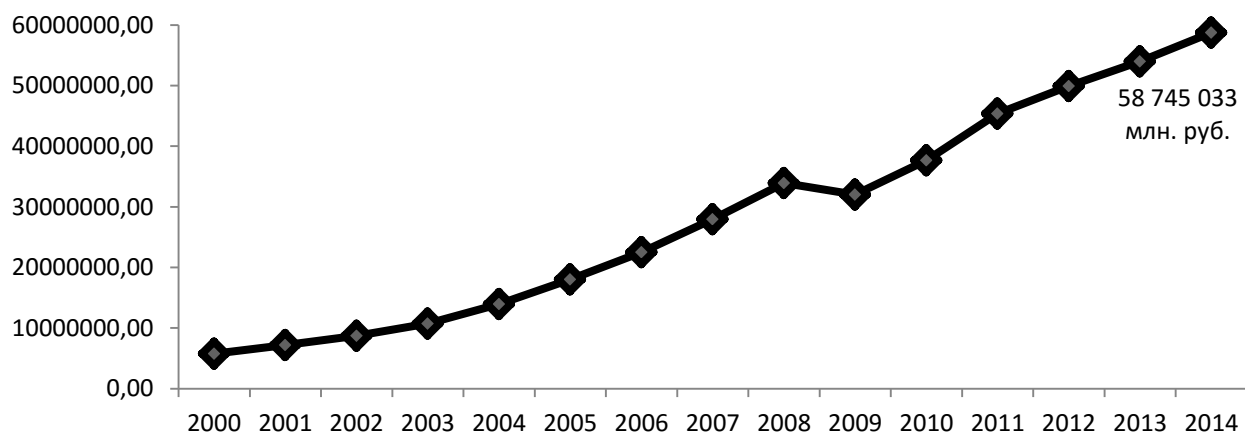


Рисунок 1.3 ВВП в России (млн руб)

Ещё одним индикатором развития регионов является уровень прямых инвестиций. С помощью данного показателя можно оценить, однородны ли суммы средств, которые направляются на развитие основных фондов регионов.

Инвестиции в экономики регионов России [млн.руб] – сумма затрат, которые были направлены на строительство, реконструкцию, модернизацию основных фондов. Данные затраты должны повышать начальную стоимость объекта, либо быть направлены на покупку транспортных средств, оборудования и т.п.

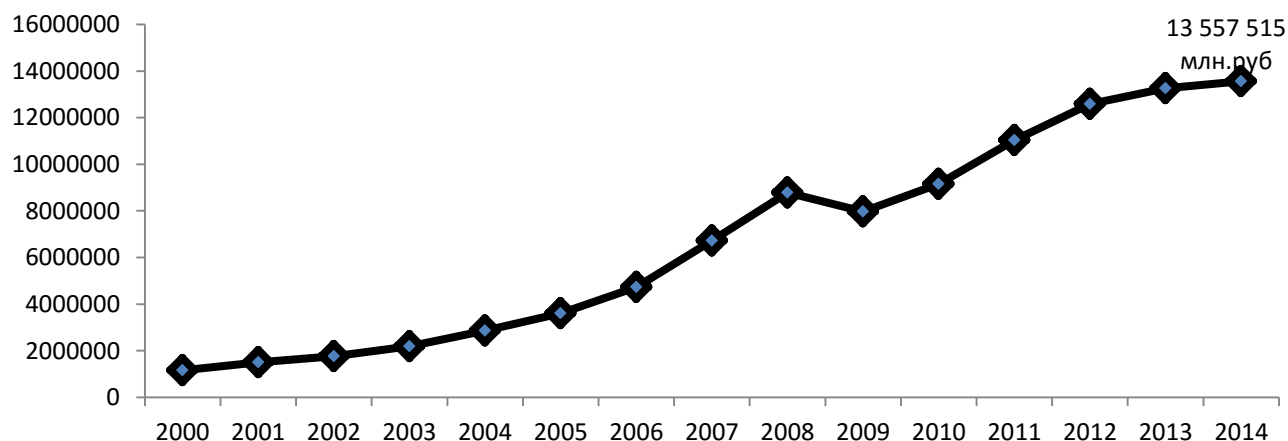


Рисунок 1.4 Российская Федерация. Инвестиции (млн.руб)

На рисунке 1.4 представлен график суммарных прямых инвестиций во все регионы России. В целом, тенденции развития уровня ВВП и уровня прямых инвестиций совпадают.

Инвестиции являются достаточно важным показателем, ввиду того, что благодаря инвестированию регионы Российской экономики имеют возможность для развития своих предприятий, увеличения объёмов производства. Благодаря инвестициям возможно совершенствование производственных и технологических процессов отечественных предприятий. Соответственно, при высоком уровне инвестиций в регионе можно сделать предположение о том, что производство совершенствуется в данном регионе.

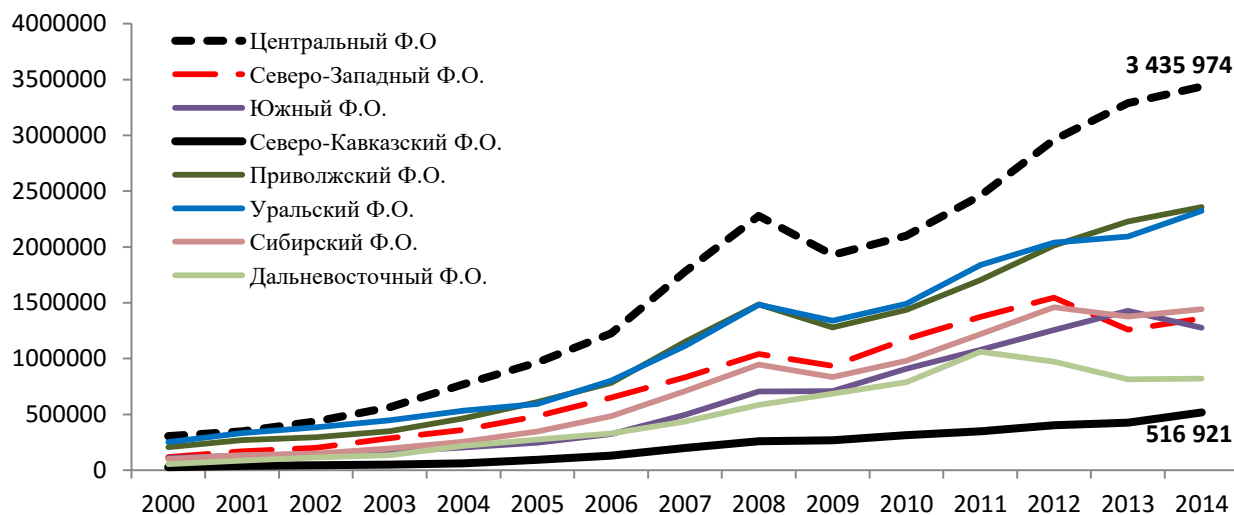


Рисунок 1.5 Инвестиции в Федеральные округа (млн.руб)

На рисунке 1.5 видно, что уровень поступающих инвестиций в регионы различен. В 2014 году разница между инвестициями в Центральный ФО и Северо-Кавказский ФО составляет чуть меньше чем 3 000 000 млн. руб.

Не представляется возможным рассматривать Федеральные округа не уделяя должного внимания на показатель занятости населения. Занятость является достаточно важной характеристикой экономики региона и страны в целом. Также занятость характеризует рынок труда региона. Уровень занятости является важной характеристикой экономического роста, т.к. при высоком уровне безработицы не возможен ни экономический рост страны, ни развитие отдельного региона.

Среднегодовая занятость в экономике регионов России [тыс.чел] – это показатель, который формируется исходя из расчета баланса трудовых ресурсов по месту работы.

К занятым относят тех лиц, которые за интересующий период:

- оплачиваемую работу (по найму);
- работу не по найму, но приносящую доход с привлечением или без наёмных рабочих;

- работу на семейном предприятии;
- отсутствовали на рабочем месте;
- были заняты в домашнем хозяйстве и производили товары и услуги для реализации.

Также, в данный показатель включают работающих иностранных граждан, которые временно или постоянно проживают на территории РФ.

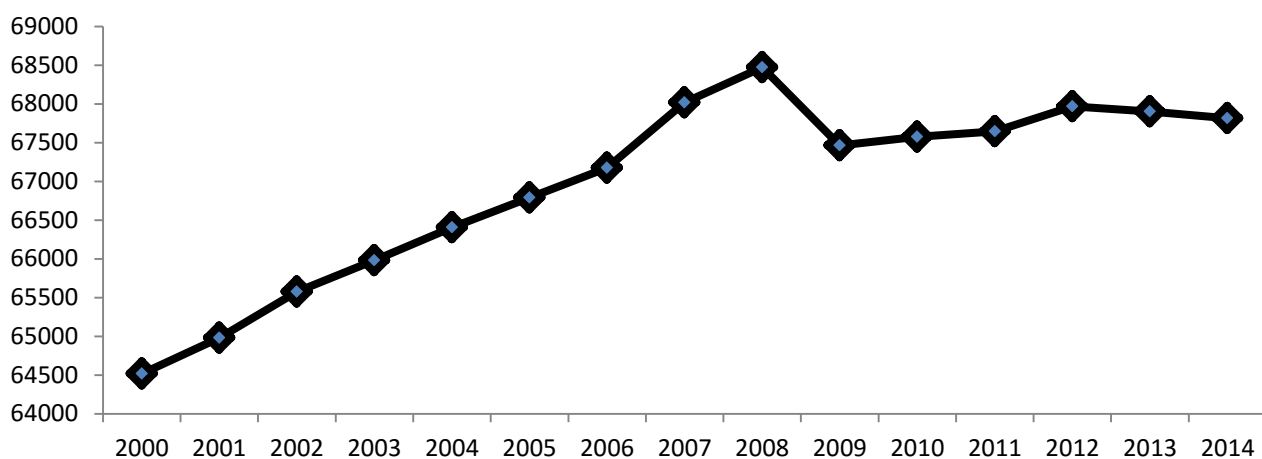


Рисунок 1.6 Россия. Среднегодовая занятость в экономике (тыс. чел)

В 2014 году уровень безработицы в Центральном ФО составил 0,7 %, а в Северо-Кавказском, в свою очередь, данный показатель был зарегистрирован на уровне 4,3%. Соответственно, уровень безработицы в Северо-Кавказском ФО более чем в 6 раз выше, чем в Центральном ФО.

Среднедушевые денежные доходы населения регионов России [руб] — сумма денежных средств, которую заработал усредненный работник за месяц. Данный показатель рассчитывается с помощью деления годовой суммы доходов на количество месяцев (12) и на среднегодовую численность населения.

Существует разница не только между количеством занятых и безработных, но и уровень оплаты труда различен. К примеру, в 2014 году разница между Центральным и Северо-Кавказским ФО составила порядка 14 тысяч рублей. В Центральном ФО средняя заработная плата составила 34 970 рублей, а в Северо-Кавказском ФО в среднем работающие получают 20 692 рублей.

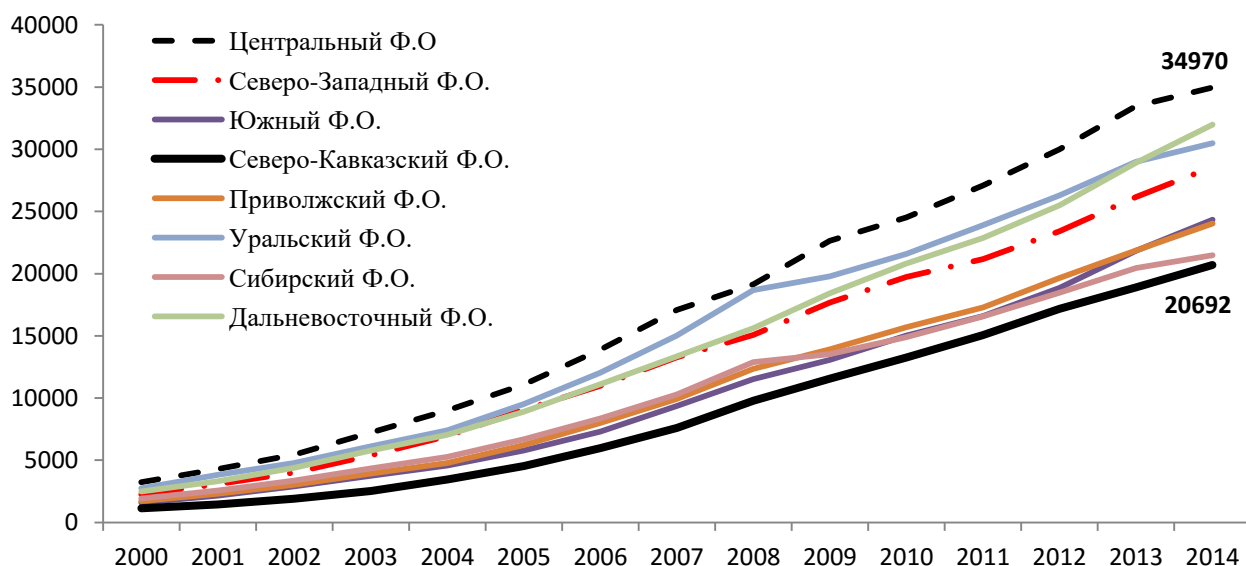


Рисунок 1.7 Различия средней заработной платы в РФ

Рассмотрев показатели ВРП, уровня инвестирования в регионы, среднюю заработную плату и уровень безработицы каждого региона, можно сделать вывод о том, что регионы развиваются не равномерно и отмечается существенная разница, к примеру, между Центральным ФО и Северо-Кавказским ФО.

1.2 Причины и возможные последствия регионального неравенства

Существуют так называемые объективные причины неравенства регионов. Эти причины были рассмотрены в 2010 году работе Н.В. Зубаревич⁴.

Одной из естественных причин является тот факт, что некоторые регионы обладают конкурентными преимуществами сразу по нескольким показателям. К примеру, Центральный ФО обладает наибольшим уровнем оплаты труда, самым высоким уровнем инвестирования и другие показатели являются более привлекательными по сравнению с остальными регионами. Как следствие, в ЦФО наблюдается высокий уровень ВРП и низкий уровень безработицы.

П. Кругманом⁵ рассматривается систематизация таких конкурентных преимуществ отдельных территорий. Выделяются факторы «первой природы» - т.е. обеспеченность природными ресурсами, которые необходимы рынку и производству. Здесь же рассматривается положение региона с точки зрения географии, транспортной доступности и т.п. К факторам «второй природы» относятся такие преимущества, которые созданы с

⁴ http://ecsocman.hse.ru/data/2015/05/25/1251201390/15-26_Zubarevich.pdf

⁵ Fujita M., Krugman P. The New Economic Geography: Past, Present and the Future // Papers in Regional Science. — Wiley-Blackwell, 2004. — Vol. 83. — P. 139—164.

помощью деятельности человека. Таковыми являются: эффект агломерации, человеческий капитал, образование, условия для развития бизнеса, инноваций.

При развитии общества и экономики преимущества первой группы факторов снижается, важнейшими и определяющими становятся факторы второй группы.

Экономические различия регионов, которые измеряются Валовым Региональным Продуктом, не так существенны для населения. Общество волнует больше социальное неравенство – различные уровни доходов населения, занятости и безработицы, образования и т.д. Социальное неравенство затормаживает рост всего государства, имеет негативное влияние на развитие, как человеческого капитала, так и различных институтов.

Одним из способов борьбы с социальным неравенством в регионах рассматривается миграция населения в наиболее благоприятные регионы. Существование данного метода находит своё подтверждение в достаточно высоком уровне миграции населения в Центральный ФО. Однако, миграция является неоднозначным процессом, с одной стороны – миграция трудовых ресурсов помогает ликвидировать дефицит рабочей силы, с помощью конкуренции работников увеличивается качество труда местного населения, увеличивается спрос на товары, которые реализуются в данном регионе. Но в то же время, отрицательное влияние миграции является достаточно важным. Ситуация на рынке рабочей силы, как правило, ухудшается и усложняется, т.к. местные жители бывают вытеснены с этого рынка ввиду миграции, которая выступает в роли дешевой рабочей силы. Мигранты оказывают существенную нагрузку на социальную инфраструктуру (сады, школы, мед. учреждения). Немаловажным является негативное отношение коренных жителей к приезжающим мигрантам. А также, большие скопления приезжих пытаются навязать культуру и традиции тому месту, в которое они приехали. Происходит коинтеграция культур, а чаще всего – приезжие не принимают местные традиции и устои и, соответственно, могут возникать межэтнические конфликты. Соответственно, пытаясь решить проблему социального неравенства регионов с помощью миграционных процессов, возникают новые вопросы и трудности.

Ещё одним фактором межрегионального неравенства внутри одной страны являются государства-соседи конкретных округов. К примеру, Соседство Дальневосточного и Сибирского ФО с Китаем не может избежать непосредственного влияния на округа данной страны с точки зрения экономики и миграции населения. Близость Скандинавских стран с Северо-Западным ФО также накладывает свой отпечаток на развитие округа.

Деятельность государства, как важнейшего политического института, должна быть направлена на уменьшение социального неравенства между субъектами одной страны. Считается, что следует вводить перераспределение финансовых ресурсов с помощью различных инструментов. Требуется помогать как населению и бизнесу отсталых регионов с помощью льгот и субсидий, так и уделять особое внимание развитым территориям для того, чтобы развивать их потенциал.

Выводы по главе I

1. Регионы России развиваются неравномерно ввиду большой территории страны и различных условий для проживания и развития бизнеса.
2. Существуют различия не только экономические, но и социальные, которые заключаются в дифференциации уровня оплаты труда, занятости и безработицы.
3. Рассматривая показатели ВРП, уровня инвестиций, уровня безработицы и среднего уровня оплаты труда, можно заметить, что Центральный ФО отличается наилучшими значениями этих показателей, а наименее - Северо-Кавказский ФО.
4. Необходима финансовая помощь регионам для обеспечения достойных уровней оплаты труда некоторым категориям граждан.
5. Требуется «выращивание» преимуществ отсталых регионов для их развития и привлечения инвестиций и капитала.
6. Неоспоримо, важным фактором развития регионов является формирования человеческого капитала, следовательно, требуется уделять особое внимание образованию.

Системы одновременных уравнений – это такое множество эконометрических уравнений, которое определяет взаимозависимости между экономическими переменными, при которой одни и те же переменные могут быть как справа, так и слева в разных уравнениях имеющейся системы.

Сами уравнения, виды связей между переменными определяются с помощью законов экономики и экономической теории. Зависимыми переменными могут быть не только экзогенные переменные, но и значения эндогенных переменных с лагом. Такие переменные, как климатические условия, социального положения, пола, возраста, так называемые, неэкономические, могут входить в уравнения только в качестве экзогенных переменных. Также экзогенными могут быть эндогенные переменные за предшествующие временные периоды.

Приведенная форма модели – это такая система, в которой все внутренние переменные выражаются через экзогенные:

В данной формуле δ_{ij} – являются коэффициентами приведенной форму модели, а u_i – это остатки для этой приведенной формы модели.

Приведенная форма выражается в виде системы независимых уравнений. Следовательно, её параметры можно оценивать с помощью обычного Метода Наименьших Квадратов.

Несомненным плюсом приведенной формы модели является то, что можно получить значения внутренних переменных через значения внешних, но отрицательным является тот факт, что в приведенной форме отсутствуют оценки взаимосвязи между внутренними переменными, а так же сами коэффициенты трудно интерпретировать с экономической точки зрения.

Введём следующие обозначения:

$$y_t = \begin{bmatrix} y_{1t} \\ \dots \\ y_{mt} \end{bmatrix}, x_t = \begin{bmatrix} x_{1t} \\ \dots \\ x_{kt} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{1t} \\ \dots \\ \varepsilon_{mt} \end{bmatrix}, B = \begin{bmatrix} \beta_{11} & \dots & \beta_{1m} \\ \vdots & \ddots & \vdots \\ \beta_{m1} & \dots & \beta_{mm} \end{bmatrix}, \Gamma = \begin{bmatrix} \gamma_{11} & \dots & \beta_{1k} \\ \vdots & \ddots & \vdots \\ \gamma_{m1} & \dots & \beta_{mk} \end{bmatrix}$$

Мы можем записать систему одновременных уравнений с помощью данных обозначений в векторно-матричной форме.

$$B * y_t + \Gamma * x_t = \varepsilon_t \quad (2.3)$$

Помимо регрессионных (поведенческих) уравнений в модели могут содержаться тождества, в виде алгебраических соотношений эндогенных переменных.

Важнейшее требование к экзогенным переменным – вектора x_t и ε_t должны быть некоррелированными в каждом t . Перечислим основные предположения⁶:

1. Математическое ожидание ошибок равно нулю, т.е. $E(\varepsilon_t) = 0$;
2. $E(\varepsilon_t \varepsilon_s') = \Sigma$ причем матрица Σ не зависит от t и положительно определена;
3. При $t \neq s$ векторы ε_t и ε_s некоррелированы;
4. Матрица B является невырожденной.

Ввиду того, что матрица B невырождена, помножим обе части уравнения на B^{-1} :

$$y_t = -B^{-1} * \Gamma * x_t + B^{-1} * \varepsilon_t = \Pi * x_t + v_t \quad (2.4)$$

В данном уравнении:

⁶ Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс: Учебник – 6-е изд., перераб. и доп. – М.: Дело, 2004. – 233

$$\begin{aligned} \Pi &= -B^{-1} * \Gamma \\ v_t &= B^{-1} * \varepsilon_t \end{aligned} \quad (2.5)$$

Выражения, записанные в соотношении (2.5), называют приведенной формой модели.

Видно, что эндогенные (внутренние) переменные и ошибки в структурной форме коррелированы, следовательно, применять для оценки метод МНК нельзя, т.к. он даёт несостоятельные и смещённые оценки для структурных коэффициентов.

Идентификация – это возможность единственным образом выразить структурные коэффициенты через приведенные коэффициенты модели.⁷ Уравнение в структурной форме идентифицируемо, если идентифицируемы все его коэффициенты. По выполнению свойства идентифицируемости структурные модели подразделяются на три следующих вида: идентифицируемые, неидентифицируемые, сверхидентифицируемые.

Модель является **идентифицируемой** в том случае, если все её структурные коэффициенты можно единственным образом выразить через коэффициенты приведенной формы. В таком случае возможно однозначно получить оценки структурных коэффициентов.

Модель является **неидентифицируемой**, если количество приведенных коэффициентов строго меньше числа структурных коэффициентов, соответственно, структурные коэффициенты невозможно оценить через коэффициенты приведенной формы. Данная модель является нерешаемой.

Модель является **сверхидентифицируемой** в том случае, если число приведенных коэффициентов строго больше количества структурных коэффициентов. Данную модель можно решить с помощью специальных методов оценки параметров. Сверхидентифицируемая модель включает в себя как минимум одно сверхидентифицируемое уравнение. Обозначим:

- число эндогенных факторов в i -м структурном уравнении = H ,
- число экзогенных факторов, не входящих в данное уравнение = D

Запишем **необходимое условие идентифицируемости** модели:

⁷ Белько И. В. Эконометрика. Практикум: Учебное пособие / И. В. Белько, Е. А. Криштапович. – Минск: Изд-во Гревцова, 2011. – С. 203

Счётное правило (условие порядка):

$D+1=H$, т.е. число эндогенных факторов в данном структурном уравнении должно быть на единицу меньше, чем число экзогенных факторов, которые не входят в это уравнение. Если это условие выполняется, следовательно, уравнение идентифицируемо;

$D+1 < H$, следовательно, уравнение **неидентифицируемо**;

$D+1 > H$, следовательно, уравнение **сверхидентифицируемо**;

Запишем **достаточное условие идентифицируемости** (условие ранга).

Для того чтобы уравнение, которое входит в систему, было идентифицируемо, необходимо и достаточно, чтобы ранг матрицы коэффициентов данной системы по отсутствующим переменным (и эндогенным, и экзогенным) был на единицу меньше, чем число внутренних переменных в системе, а также определитель данной матрицы должен быть не равен нулю.

В зависимости от того, какого вида СОУ, а также от того, какой вид связей установлен между её переменными, требуются различные способы формирования оценок структурных коэффициентов.

Существуют различные методы оценивания коэффициентов структурной модели:

- Косвенный метод наименьших квадратов (КМНК);
- Двухшаговый метод наименьших квадратов (ДМНК);
- Трёхшаговый метод наименьших квадратов (ТМНК);
- Метод максимального правдоподобия;
- Метод инструментальных переменных.

2.1.1 Методы оценивания*Косвенный МНК*

Данный метод целесообразно применять, когда структурная модель точно идентифицируема. Рассмотрим процедуру применения КМНК:

1. Приведение структурной модели в приведенную форму;
2. Оценка приведенных коэффициентов для всех приведенных уравнений данной системы с помощью традиционного МНК;

3. Преобразование приведенных коэффициентов в параметры структурной модели.

Таким образом, оценивается приведенная форма модели, далее следует решить систему, в которой вместо элементов матрицы Π записываются её оценки.

Рассмотрим более подробно процедуру оценки с практической точки зрения. Запишем исходные уравнения. Для простоты будем использовать два уравнения с двумя эндогенными переменными. Данное упрощение не влияет на качество полученных результатов и возможность их применения для систем другой размерности.

$$\begin{aligned} Y_1 &= a_1 + \beta_1 * X_1 + \gamma_1 * Y_2 + \varepsilon_1 \\ Y_2 &= a_2 + \beta_2 * X_2 + \gamma_2 * Y_1 + \varepsilon_2 \end{aligned} \quad (2.6)$$

Далее, решим уравнения (2.6) относительно Y_1 и Y_2 .

$$\begin{aligned} Y_1 &= \frac{a_1 + \gamma_1 * a_2}{1 - \gamma_1 * \gamma_2} + \frac{\beta_1}{1 - \gamma_1 * \gamma_2} * X_1 + \frac{\gamma_1 * \beta_2}{1 - \gamma_1 * \gamma_2} * Y_2 + \frac{\gamma_1 * \varepsilon_2 + \varepsilon_1}{1 - \gamma_1 * \gamma_2} \\ Y_2 &= \frac{a_2 + \gamma_2 * a_1}{1 - \gamma_1 * \gamma_2} + \frac{\gamma_2 * \beta_1}{1 - \gamma_1 * \gamma_2} * X_2 + \frac{\beta_2}{1 - \gamma_1 * \gamma_2} * Y_1 + \frac{\gamma_2 * \varepsilon_1 + \varepsilon_2}{1 - \gamma_1 * \gamma_2} \end{aligned} \quad (2.7)$$

Введём обозначения для упрощения:

$$\begin{aligned} a_1 &= \frac{\alpha_1 + \gamma_1 * \alpha_2}{1 - \gamma_1 * \gamma_2}, a_2 = \frac{\alpha_2 + \gamma_2 * \alpha_1}{1 - \gamma_1 * \gamma_2}, b_1 = \frac{\beta_1}{1 - \gamma_1 * \gamma_2}, b_2 = \frac{\gamma_2 * \beta_1}{1 - \gamma_1 * \gamma_2}, \\ c_1 &= \frac{\gamma_1 * \beta_2}{1 - \gamma_1 * \gamma_2}, c_2 = \frac{\beta_2}{1 - \gamma_1 * \gamma_2}, v_1 = \frac{\beta_2}{1 - \gamma_1 * \gamma_2}, v_2 = \frac{\gamma_2 * \varepsilon_1 + \varepsilon_2}{1 - \gamma_1 * \gamma_2}. \end{aligned} \quad (2.8)$$

Запишем модель (2.7) в более краткой форме с учетом введенных обозначений в формулах (2.8).

$$\begin{aligned} Y_1 &= a_1 + b_1 * X_1 + c_1 * Y_2 + v_1 \\ Y_2 &= a_2 + b_2 * X_2 + c_2 * Y_1 + v_2 \end{aligned} \quad (2.9)$$

Введём следующее упрощение: Пусть все Y имеют математическое ожидание = 0, т.е. переменные отцентрированы. Данное упрощение является несущественным для практических выводов.

Теперь, применив к (2.9) традиционный МНК, можно получить оценки для параметров b и c .

$$\begin{aligned}\hat{b}_1 &= \frac{\langle X_2 X_2 \rangle \langle X_1 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_1 \rangle}{\langle X_1 X_1 \rangle \langle X_2 X_2 \rangle - \langle X_1 X_2 \rangle^2}, \hat{c}_1 = \frac{\langle X_1 X_1 \rangle \langle X_2 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_1 Y_1 \rangle}{\langle X_1 X_1 \rangle \langle X_2 X_2 \rangle - \langle X_1 X_2 \rangle^2} \\ \hat{b}_2 &= \frac{\langle X_2 X_2 \rangle \langle X_1 Y_2 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_2 \rangle}{\langle X_1 X_1 \rangle \langle X_2 X_2 \rangle - \langle X_1 X_2 \rangle^2}, \hat{c}_2 = \frac{\langle X_1 X_1 \rangle \langle X_2 Y_2 \rangle - \langle X_1 X_2 \rangle \langle X_1 Y_2 \rangle}{\langle X_1 X_1 \rangle \langle X_2 X_2 \rangle - \langle X_1 X_2 \rangle^2}\end{aligned}\quad (2.10)$$

$$\text{Где: } \langle X_i X_j \rangle = \sum_{i=1}^n x_i x_j, \quad \langle Y_i Y_j \rangle = \sum_{i=1}^n y_i y_j, \quad \langle X_i Y_j \rangle = \sum_{i=1}^n x_i y_j,$$

Далее возможно выразить через оценки, полученные в (2.10), оценки исходных параметров.

$$\begin{aligned}\hat{b}_1 &= \frac{\langle X_1 Y_2 \rangle \langle X_2 Y_2 \rangle - \langle X_2 Y_1 \rangle \langle X_1 Y_2 \rangle}{\langle X_1 X_1 \rangle \langle X_2 Y_2 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_1 \rangle}, \hat{b}_2 = \frac{\langle X_1 Y_1 \rangle \langle X_2 Y_2 \rangle - \langle X_2 Y_1 \rangle \langle X_1 Y_2 \rangle}{\langle X_2 X_2 \rangle \langle X_1 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_1 \rangle} \\ \hat{\gamma}_1 &= \frac{\langle X_1 X_1 \rangle \langle X_2 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_1 Y_1 \rangle}{\langle X_2 X_2 \rangle \langle X_2 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_1 \rangle}, \hat{\gamma}_2 = \frac{\langle X_2 X_2 \rangle \langle X_1 Y_2 \rangle - \langle X_1 X_2 \rangle \langle X_1 Y_2 \rangle}{\langle X_2 X_2 \rangle \langle X_2 Y_1 \rangle - \langle X_1 X_2 \rangle \langle X_2 Y_1 \rangle}\end{aligned}\quad (2.11)$$

Оценки (2.11), которые были получены с помощью косвенного метода наименьших квадратов, являются состоятельными, в отличие от оценок, полученных при применении традиционного МНК.

Двухшаговый МНК

Данный метод может применяться, если модель сверхидентифицируема. В данном случае не корректно применять косвенный МНК, ввиду того что он не может дать однозначно определенных оценок параметров структурной модели.

Основной идеей ДМНК является получение для сверхидентифицируемого уравнения теоретических значений внутренних переменных из правой части уравнения на основе приведенной формы модели. Затем их подставляют вместо фактических значений, применяется традиционный МНК к структурной форме сверхидентифицируемого уравнения. Данный метод двухшаговый, т.к. мы используем два раза МНК, первый раз – когда определяется приведенная форма модели и происходит поиск оценок теоретических значений внутренней переменной. Второй раз МНК применяется, когда оценивается структурное сверхидентифицируемое уравнение, чтобы определить структурные коэффициенты модели по имеющимся теоретическим значениям внутренних переменных.

На практике может быть два типа сверхидентифицируемой структурной модели:

1. Все уравнения системы сверхидентифицируемы;
2. Система содержит и точно идентифицируемые уравнения.

Рассмотрим применение процедуры двухшагового МНК более подробно.

Запишем систему уравнений в векторно-матричном виде:

$$y_1 = Y_1 * \beta_1 + X_1 * \gamma_1 + \varepsilon_1 \quad (2.12)$$

Где: $y_1 = \begin{bmatrix} y_{11} \\ y_{12} \\ \dots \\ y_{1n} \end{bmatrix}$, $Y_1 = \begin{bmatrix} y_{11} & \dots & y_{q1} \\ y_{12} & \dots & y_{q2} \\ \dots & \dots & \dots \\ y_{1n} & \dots & y_{qn} \end{bmatrix}$, $X_1 = \begin{bmatrix} x_{11} & \dots & x_{p1} \\ x_{12} & \dots & x_{p2} \\ \dots & \dots & \dots \\ x_{1n} & \dots & x_{pn} \end{bmatrix}$, $\beta_1 = \begin{bmatrix} \beta_{11} \\ \beta_{12} \\ \dots \\ \beta_{1q} \end{bmatrix}$, $\gamma_1 = \begin{bmatrix} \gamma_{11} \\ \gamma_{12} \\ \dots \\ \gamma_{1p} \end{bmatrix}$, $\varepsilon_1 = \begin{bmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \dots \\ \varepsilon_{1n} \end{bmatrix}$.

Если к такой модели, в которой элементы Y_1 коррелируют с остатками ε_1 применить традиционный метод наименьших квадратов, то будут получены смещенные и несостоятельные оценки.

Проведём регрессию каждого столбца матрицы Y_1 на все внешние переменные. Таким образом, рассмотрим регрессию:

$$Y_1 = X * \Pi_1 + V_1 \quad (2.13)$$

Матрица Π_1 — это $k \times (q-1)$ матрица, содержащая коэффициенты приведенной формы.

Далее, требуется построить прогнозное значение $\hat{Y}_1 = X * \hat{\Pi}_1$, где $\hat{\Pi}_1 = (X'X)^{-1} * X'Y_1$.

Затем, осуществим регрессию каждого столбца матрицы Y_1 ($Y_1 = X * \Pi_1 + V_1$), заменив Y_1 на \hat{Y}_1 . Таким образом, мы построили МНК-оценки структурных параметров β_1 и γ_1 .

$$y_1 = \hat{Y}_1 * \beta_1 + X_1 * \gamma_1 + \varepsilon_1 \quad (2.14)$$

Перечислим основные выводы, связанные с применением Двухшагового МНК⁸:

1. Если выполняются необходимое и достаточное условия идентификации, и уравнения являются точно идентифицируемыми, то оценка, полученная при применении ДМНК, совпадает с оценкой, полученной при оценивании КМНК.
2. Полученная оценка при применении ДМНК совпадает с оценкой метода инструментальных переменных, если в качестве инструментов используются \hat{Y}_1 и X_1 .
3. Если инструментальными переменными для Y_1 взяты любые линейные комбинации столбцов матрицы X , то матрица ковариаций полученной оценки больше или равна, чем матрица ковариаций, полученной при применении ДМНК.

⁸ Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс: Учебник – 6-е изд., перераб. и доп. – М.: Дело, 2004. – С. 239

Метод максимального правдоподобия

Данный метод позволяет получить асимптотически несмещенные и эффективные оценки параметров распределения, имеющие нормальный закон распределения.

Рассматривается вектор $X = (X_1, X_2, \dots, X_n)$, с плотностью распределения $(x; \theta)$, где θ – неизвестный параметр. Функция правдоподобия – это случайная величина L , где $L = L(x; \theta) = p(x; \theta)$. Оценкой максимального правдоподобия является следующая величина $\hat{\theta} = \hat{\theta}_{ML}$, которая максимизирует функцию L . Таким образом, оценка максимального правдоподобия, это такая функция $\hat{\theta} = \hat{\theta}(x)$, которая выполняет условие: $L(x; \hat{\theta}(x)) = \max_{\theta} L(x, \theta)$. Необходимым условием является равенство нулю частной производной функции $L(x, \theta)$ по переменной θ , т.е. $\frac{dL(x; \theta)}{d\theta} = 0$. Оценки максимального правдоподобия состоятельны, а также асимптотически эффективны. Но они могут быть смещенными.

Метод инструментальных переменных

Данный метод применяется для точно идентифицируемых СОУ. Предопределенные факторы структурных уравнений коррелируют со случайными остатками, следовательно, традиционный (обычный) МНК неприменим, т.к. МНК-оценки не являются состоятельными. Требуется введение новых переменных для улучшения данных оценок. Для структурного уравнения рассматривается вектор наблюдений той же самой эндогенной переменной и вектор значений соответствующей компоненты случайных остатков. Вектор инструментальных переменных – случайный вектор, имеющий невырожденную ковариационную матрицу, не коррелированный с вектором случайных остатков и имеет невырожденную ковариационную матрицу с внутренними и предопределенными переменными, входящими в j -е структурное уравнение. Далее, на основе такого вектора строят состоятельную МИП-оценку вектора параметров j -го структурного уравнения. Точность этой оценки зависит от корреляции с эндогенными (внутренними) и предопределенными переменными, которые входят в j -е структурное уравнение. Точность тем выше, чем больше значение корреляции.

Рассмотрим в отдельности два случая, когда система идентифицируема и когда она не идентифицируема.

1 случай. Система идентифицируема.

Применяя экзогенные переменные в качестве инструментальных переменных, получаются оценки, которые совпадают с оценками, полученными при применении процедуры КМНК.⁹ Следовательно, делается вывод о том, что косвенный МНК – это частный случай метода инструментальных переменных.

Достаточно часто в практических исследованиях метод инструментальных переменных (МИП) применяется в виде Двухшагового МНК. Таким образом, в качестве инструментальных переменных используют объясненные значения этих переменных, которые получаются в результате оценки приведенной формы. Далее, эти значения подставляют в правую часть структурной формы уравнения, которая была рассмотрена (2.9).

Следовательно, возможно сформулировать ещё один вывод. Если система идентифицируема, и количество внешних (экзогенных) переменных X равняется количеству внутренних (эндогенных) переменных Y , то оценки Двухшагового МНК совпадают с оценками косвенного МНК.

2 случай. Система неидентифицируема.

В таком случае МИП используют, так называемые, «внешние» переменные. Если у нас имеется число инструментальных переменных l , и оно избыточно, а также мы можем использовать любые наборы данных инструментальных переменных. Таким образом, двухшаговый МНК – это наилучший возможный выбор.

Пусть $\{Z\}$ – набор инструментальных переменных (экзогенных и внешних). \hat{Y}_i – Это проекции эндогенных факторов на пространство Z . Запишем соответствующие регрессии, полученные с помощью традиционного МНК.

$$\hat{Y}_i = a_i + \sum_{j=1}^l b_{ij} * Z_j \quad (2.15)$$

Следовательно, переменные \hat{Y}_i являются линейными комбинациями инструментальных переменных, которые достаточно сильно коррелируют с Y_i . Такую замену \hat{Y}_i на Y_i называют «очищением» внутренней переменной.

⁹ Кремер Н. Ш., Путко Б. А. Эконометрика: Учебник для вузов / Под ред. проф. Н. Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. – С. 233

2.1.2 Инструментальные переменные. Тест Хаусмана

Инструментальные переменные:

При существовании корреляции между ошибками модели, а также независимыми переменными, МНК - оценки бывают несостоятельные и смещенные. В качестве решения данной проблемы можно использовать другие независимые переменные. Переменные такого рода называются инструментальными. Важно, чтобы данные переменные обладали следующими свойствами¹⁰:

1. Такие новые независимые переменные должны «хорошо коррелировать» с начальными независимыми;
2. Новые переменные не должны коррелировать с ошибками.

Опишем исходную модель с помощью следующего равенства:

$$y = X * \beta + \varepsilon$$

Где: y – вектор-столбец зависимых переменных, размерности $n \times 1$;

ε – вектор-столбец ошибок, размерности $n \times 1$;

β - $k \times 1$ вектор-столбец параметров модели.

Пусть матрица Z , с размерностью $n \times k$ – матрица инструментальных переменных. Запишем оценку параметров β через инструментальные переменные:

$$\hat{\beta}_{IV} = (Z' * X)^{-1} * Z' * y \quad (2.16)$$

Подставим первоначальное выражение для y в формулу (2.16):

$$\hat{\beta}_{IV} = (Z' * X)^{-1} * Z' * (X * \beta + \varepsilon) = \beta + (Z' * X)^{-1} * Z' * \varepsilon = \beta + \left(\frac{1}{n} * Z' * X\right)^{-1} * \frac{1}{n} * Z' * \varepsilon \quad (2.17)$$

Полученная формула (2.17) выражает требование «хорошей корреляции» между элементами матрицы X и Z : последовательность матриц $\frac{1}{n} * Z' * X$ сходится по вероятности при $n \rightarrow \infty$ к некоторой невырожденной матрице. Ввиду того, что отсутствует корреляция между элементами матрицы Z и ошибками, $\frac{1}{n} * Z' * \varepsilon \rightarrow 0$ по вероятности.

¹⁰ Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс: Учебник – 6-е изд., перераб. и доп. – М.: Дело, 2004. – С. 212

Аналогично, условие того, что матрица Z и ошибки не коррелированы, можно записать в виде такого условия: $p * \lim_{n \rightarrow \infty} \frac{1}{n} * Z' * \varepsilon = 0$

Теперь можно сделать выводы о том, что оценка $\hat{\beta}_{IV}$ – состоятельна. Т.к. оценка $\hat{\beta}_{IV}$ зависит от Z , следовательно, не является эффективной.

Самым важным вопросом является то, как именно находить требующиеся инструментальные переменные. Явного решения этого вопроса для всех случаев нет. Может быть ситуация, когда существует несколько инструментальных переменных, а может их не быть вовсе. Чаще всего необходимо пользоваться инструментальными переменными в том случае, если есть ошибки в измерениях.

Ошибки в измерениях зависимой переменной

Пусть у нас есть модель $y = X * \beta + \varepsilon$, а вектор y оценивается с ошибкой (u , математическое ожидание равно 0), т.е. мы можем наблюдать вектор $y^* = y + u$, u – не зависят от X и ε . Следовательно, построение МНК-оценок аналогично регрессии:

$$y^* = X * \beta + (\varepsilon + u) \quad (2.18)$$

Отсюда видно, что $\hat{\beta}$ – несмещенная и состоятельная, ввиду того, что математическое ожидание величины $(\varepsilon + u)$ равно нулю. Также ковариация между X и $(\varepsilon + u)$ отсутствует. Наличие ошибки u приводит к тому, что дисперсия регрессии увеличивается:

$$V(\varepsilon + u) = \sigma_{\varepsilon}^2 + \sigma_u^2 \quad (2.19)$$

Ошибки в измерениях независимой переменной

Пусть есть также исходная модель, $y = X * \beta + \varepsilon$. Также наблюдается матрица $X^* = X + V$, матрица V – с математическим ожидаем = 0, и также не зависит от ошибок ε . Следовательно, существует регрессия:

$$y = X^* * \beta + (\varepsilon - V * \beta) = X^* * \beta + \varepsilon^* \quad (2.20)$$

В данном случае заметна сильная корреляция между регрессорами и ошибками. Рассмотрим математическое ожидание следующих величин:

$$E(X^{*'} \varepsilon^*) = E[(X^{*'} + V')(\varepsilon - V * \beta)] = -E(V'V) * \beta \quad (2.21)$$

Таким образом, МНК-оценки являются смещенными и несостоятельными.

Не обязательно количество исходных регрессоров должно быть равно количеству инструментальных переменных. Достаточным является условие, чтобы число

инструментальных переменных было не меньше, чем количество независимых переменных.

Тест Хаусмана

Для определения, требуется ли внедрение инструментальных переменных для того, чтобы оценить модель достаточно протестировать гипотезу:

$$H_0 : p \cdot \lim_{n \rightarrow \infty} \frac{1}{n} * X' * \varepsilon = 0$$

гипотеза H_1 выглядит следующим образом:

$$H_1 : p \cdot \lim_{n \rightarrow \infty} \frac{1}{n} * X' * \varepsilon \neq 0$$

Требуется считать, что имеется МНК-оценка $\hat{\beta}_1 = \hat{\beta}_{OLS}$, а также $\hat{\beta}_2 = \hat{\beta}_{IV}$, полученная при использовании инструментальных переменных. Если гипотеза H_0 верна, то оценка $\hat{\beta}_1$ – состоятельна и эффективна. Если верна H_1 , то данная оценка несостоятельна. $\hat{\beta}_2$ состоятельна при обеих гипотезах. Следовательно, разность $(\hat{\beta}_2 - \hat{\beta}_1)$ стремится к нулю. Следует пронормировать эту разность и она будет совпадать асимптотически с известным распределением.

В 1978 году Джерри Хаусман (Jerry Hausman) доказал, что данная разность с нормировкой V , асимптотически имеют распределение χ^2 (хи-квадрат) с k -степенями свободы, следовательно:

$$V(\hat{\beta}_2 - \hat{\beta}_1) = V(\hat{\beta}_2) - V(\hat{\beta}_1) \text{ и } (\hat{\beta}_2 - \hat{\beta}_1)' (V(\hat{\beta}_2) - V(\hat{\beta}_1))^{-1} * (\hat{\beta}_2 - \hat{\beta}_1) - \text{асимптотически}$$

имеет хи-квадрат распределение с k -степенями свободы.

2.2. Построение и оценивание СОУ

Рассмотрим следующую систему уравнений:

$$\begin{cases} U_t = a_1 + b_{11} * P_t + \varepsilon_1 \\ Y_t = a_2 + b_{21} * N_t + c_{31} * U_t + \varepsilon_2 \\ I_t = a_3 + b_{31} * Y_t + b_{32} * P_t + \varepsilon_3 \end{cases}$$

Где:

$U_t = a_1 + b_{11} * P_t + \varepsilon_1$ – кривая Филлипса

$Y_t = a_2 + b_{21} * N_t + c_{31} * U_t + \varepsilon_2$ – производственная функция

$I_t = a_3 + b_{31} * Y_t + b_{32} * P_t + \varepsilon_3$ – функция инвестиций.

Кривая Филлипса – кривая, отражающая обратную зависимость между темпом инфляции и безработицы. Важным является тот факт, что одновременно не может быть высоким показатель инфляции и безработицы. Таким образом, при увеличении уровня инфляции снижается уровень безработицы, а также при снижении уровня инфляции, увеличивается уровень безработицы.

Производственная функция – функция, выражающая зависимость между величиной выпуска и факторами производства.

Функция инвестиций – отражает валовые инвестиционные расходы всех субъектов экономики.

Построим систему на основе представленных в предыдущем параграфе данных, с помощью программного продукта Gretl для Центрального федерального округа.

Система уравнений, Система 1 (Центральный ФО)

Уравнение 1:

Метод оценивания: Двухшаговый МНК, использованы наблюдения 2000-2013 (T = 14)

Зависимая переменная: **unempl_1**

Инструменты: **const, CPI_1, EAP_1.**

	Коэффициент	Ст. ошибка	z	P-значение
const	-15,6140	6,48170	-2,409	0,0160 **
CPI_1	0,181851	0,0582198	3,124	0,0018 ***

Среднее зависимой переменной = 4,614286

Стандартное отклонение зависимой переменной = 1,301985

Сумма квадратов остатков = 12,15485

Стандартная ошибка модели = 1,006431

R-квадрат = 0,448438

Исправленный R-квадрат = 0,402474

Уравнение 2:

Метод оценивания: Двухшаговый МНК, использованы наблюдения 2000-2013 (T = 14)

Зависимая переменная: **GRP_1**

Инструменты: **const, CPI_1, EAP_1.**

	Коэффициент	Ст. ошибка	z	P-значение
const	-2,10541e+08	3,07512e+07	-6,847	7,56e-012 ***
EAP_1	3,30064e+06	412364	8,004	1,20e-015 ***
unempl_1	-567222	705341	-0,8042	0,4213

Среднее зависимой переменной = 9257479

Стандартное отклонение зависимой переменной = 5893468

Сумма квадратов остатков = 1,20e+13

Стандартная ошибка модели = 1044341

R-квадрат = 0,973439

Исправленный R-квадрат = 0,968609

Уравнение 3:

Метод оценивания: Двухшаговый МНК, использованы наблюдения 2000-2013 (T = 14)

Зависимая переменная: **inv_1**

Инструменты: **const, CPI_1, EAP_1**.

	Коэффициент	Ст. ошибка	z	P-значение
const	-1,36196e+06	963379	-1,414	0,1574
GRP_1	0,176085	0,00681107	25,85	2,26e-147 ***
CPI_1	11334,7	8253,17	1,373	0,1696

Среднее зависимой переменной = 1 528 961

Стандартное отклонение зависимой переменной = 1 004 768

Сумма квадратов остатков = 1,15e+11

Стандартная ошибка модели = 102 461,3

R-квадрат = 0,991202

Исправленный R-квадрат = 0,989602

Общая матрица ковариации (cross-equation VCV) (корреляции выше главной диагонали):

0,86820	(0,191)	(-0,294)
1,6439e+005	8,5694e+011	(-0,225)
-24854	-1,8937e+010	8,2487e+009

Логарифм определителя = 50,0083

Тест Бриша-Пэгона (Breusch-Pagan) для диагональной ковариационной матрицы:

Хи-квадрат(3) = 2,42635 [0,4887]

Рассмотрим полученные результаты. Можно построить следующую систему:

$$\begin{cases} U_t = -15.61 + 0.18 * P_t + \varepsilon_1 \\ Y_t = -0.000000021 + 0.0000033 * N_t - 567222 * U_t + \varepsilon_2 \\ I_t = -0.000001362 + 0.17 * Y_t + 1133417 * P_t + \varepsilon_3 \end{cases}$$

В первом уравнении и константа, и переменная P_t являются значимыми. $R^2 = 44.8\%$.

Во втором уравнении значимая константа и переменная N являются значимыми, а

переменная U_t не значима. $R^2 = 97,3\%$. В последнем уравнении константа и P_t являются

незначимыми, а Y_t значима. $R^2 = 99,12\%$.

Проведённый тест Бриша-Пэгона показывает, что с вероятностью 48.87% отвергается

нулевая гипотеза. Нулевой гипотезой в данном тесте является наличие гетероскедастичности, т.е. что все коэффициенты при переменных равны нулю.

Далее, рассмотрим полученные системы для остальных семи федеральных округов. Расчёты для этих систем представлены в приложениях 1-7.

1. **Северо-Западный ФО**

$$\begin{cases} U_t = -17.1715 + 0.206047 * P_t + \varepsilon_1 \\ Y_t = -0.00000013 + 296533 * N_t - 827708 * U_t + \varepsilon_2 \\ I_t = -0.00000164 + 0.3194 * Y_t + 113460.1 * P_t + \varepsilon_3 \end{cases}$$

2. **Южный ФО**

$$\begin{cases} U_t = -33.3585 + 0.380486 * P_t + \varepsilon_1 \\ Y_t = 80355.2 + 80355.2 * N_t - 501218 * U_t + \varepsilon_2 \\ I_t = -0.000002096 + 0.451111 * Y_t + 18106.7 * P_t + \varepsilon_3 \end{cases}$$

3. **Северо-Кавказский ФО**

$$\begin{cases} U_t = -26.9429 + 0.399043 * P_t + \varepsilon_1 \\ Y_t = -625023 + 58508.3 * N_t - 141367 * U_t + \varepsilon_2 \\ I_t = -397690 + 0.359302 * Y_t + 3389.36 * P_t + \varepsilon_3 \end{cases}$$

4. **Приволжский ФО**

$$\begin{cases} U_t = -16.1413 + 0.210655 * P_t + \varepsilon_1 \\ Y_t = -0.0000000675 + 0.00000118 * N_t - 949138 * U_t + \varepsilon_2 \\ I_t = -0.00000149 + 0.284225 * Y_t + 12052 * P_t + \varepsilon_3 \end{cases}$$

5. **Уральский ФО**

$$\begin{cases} U_t = -16.8621 + 0.215216 * P_t + \varepsilon_1 \\ Y_t = -0.0000000475 + 839614 * N_t - 777384 * U_t + \varepsilon_2 \\ I_t = -0.0000031240 + 0.341534 * Y_t + 25662.7 * P_t + \varepsilon_3 \end{cases}$$

6. **Сибирский ФО**

$$\begin{cases} U_t = -23.5457 + 0.294602 * P_t + \varepsilon_1 \\ Y_t = -0.0000098049 + 308062 * N_t - 812110 * U_t + \varepsilon_2 \\ I_t = 0.00000255 + 0.340392 * Y_t + 20176.1 * P_t + \varepsilon_3 \end{cases}$$

7. **Дальневосточный ФО**

$$\begin{cases} U_t = -24.2643 + 0.292187 * P_t + \varepsilon_1 \\ Y_t = -0.0000002816 + 473823 * N_t - 320581 * U_t + \varepsilon_2 \\ I_t = -226135 + 0.390255 * Y_t + 1491.99 * P_t + \varepsilon_3 \end{cases}$$

Ниже представлена сводная таблица, в которой можно увидеть, какие переменные, в каких уравнениях и в каких системах значимы. Можно заметить, что только в Приволжском ФО абсолютно все переменные значимы на 1% и 5% уровнях значимости.

Таблица 2.1. Значимость переменных в уравнениях систем

№	Наименование ФО	1 уравнение		2 уравнение			3 уравнение		
		const	CPI	const	EAP	unempl	const	GRP	CPI
1	Центральный	**	***	***	***			***	
2	Северо-Западный	**	***					***	
3	Южный	**	***		***	***	**	***	**
4	Северо-Кавказский		**				***	***	***
5	Приволжский	**	***	***	***	***	**	***	**
6	Уральский	**	***	**	***	*		***	
7	Сибирский	***	***			**		***	
8	Дальневосточный	***	***	***	***	***		***	

Несмотря на то, что все федеральные округа по своей сути однородны, но даже исходя из построенных уравнений и коэффициентов становится видно, что развиваются федеральные округа по-разному. Можно заметить, что Центральный федеральный округ и Северо-Западный не возможно сопоставлять и сравнивать с такими федеральными округами как Южный и Уральский, т.к. разница в коэффициентах при переменных, а также при константе различаются в десятки раз.

Полученные результаты свидетельствуют о существенных различиях в процессе развития и существования округов. Аналогично, для наглядности можно представить показатели на следующем графике (рисунок 2.1). Объем шаров характеризует объем ВРП в 2013 году. По оси «Х» расположен уровень безработицы, по оси «У» - уровень экономической активности населения. Можно заметить, что от общей группы регионов находятся в отдалении Южный и Северо-Кавказский округа.

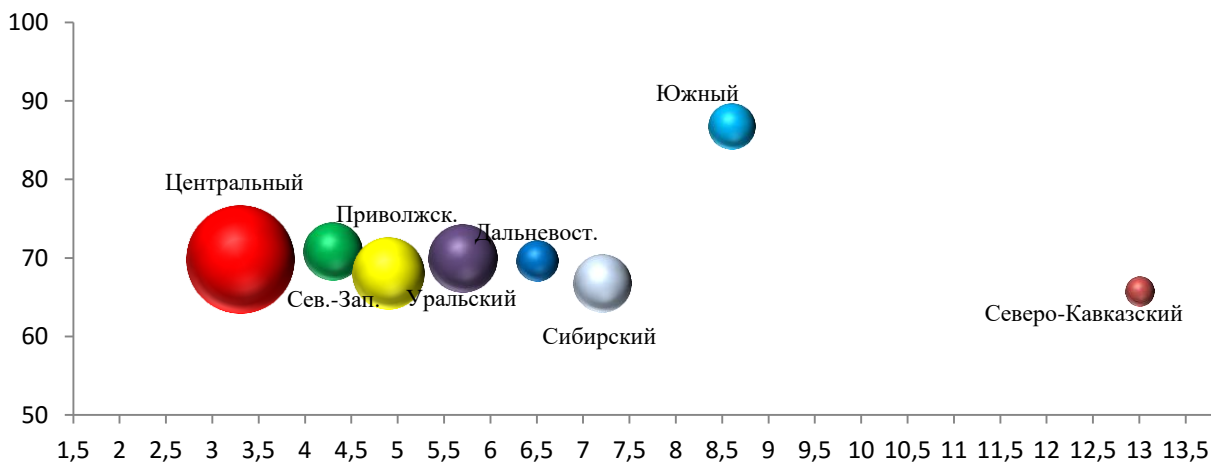


Рисунок 2.1 Диаграмма федеральных округов по уровню ВРП, безработицы и экономической активности

2.3 Модели панельных данных

2.3.1 Базовая модель панельных данных

Панельные данные представляют собой двумерный массив данных, в котором одно измерение является «пространственным» (по показателям), а другое является «временным» (по моментам времени). Следовательно, панельные данные обозначены двумя индексами. Такие массивы формируются в процессе изучения большого числа объектов за определенный период времени.

Если сравнивать панельные данные с пространственными или временными рядами данных, то можно выделить важное преимущество, которое присуще панельным данным. Панельные данные позволяют анализировать и вычленять изменения на индивидуальном уровне каждого объекта. Неоспоримым плюсом также является то, что в панельных данных используется большое количество статистических наблюдений, что позволяет увеличить число степеней свободы и уменьшить зависимость факторов за счёт того, что учитываются индивидуальные особенности зависимых переменных. Следовательно, получаются оценки, являющиеся более эффективными. Также, панельные данные позволяют учитывать влияния факторов при наличии лагов. Существуют различные виды структур панельных данных. Рассмотрим основные:

1. Сбалансированные – если на протяжении всего периода экономические показатели не меняются (со временем не появляются новые и не исчезают имеющиеся).
2. Несбалансированные – такие структуры, в которых отсутствуют некоторые экономические показатели в определенные периоды.
3. Псевдопанельные данные – возникают в том случае, если в разные моменты времени наблюдаются разные экономические показатели.

Ввиду того, что такой вид данных содержит статистику по одним единицам за различные периоды времени, то может возникнуть проблема автокорреляции, соответственно, требуются особые методы оценивания, подходящие именно под такой тип данных.

Регрессионная модель панельных данных в общем виде записывается следующим образом:

$$x_{it} = Z_{it} * a_{it} + \varepsilon_{it} \quad i = 1 \dots N, \quad t = 1 \dots T \quad (2.22)$$

Где:

i – это индекс экономического объекта (фирмы, округа, страны и т.п.),

t – это момент времени,

a_{it} – это коэффициенты вектора зависимых переменных Z_{it} в момент времени t для объекта i .

Недостатком данной модели является то, что она слишком общая и не подходит для сложных, специфических моделей, которые бы описывали реальные экономические процессы. Данную модель невозможно оценить, но можно вычленить специфические факторы, относящиеся в определенный момент времени к определенному экономическому объекту, соответственно, можно оценить индивидуальные (специфические) особенности.

$$x_{it} = Z_{it} * a + \gamma_t + f_i + \varepsilon_{it} \quad (2.23)$$

Где:

Z_{it} – это n -мерный вектор зависимых переменных без константы. Следовательно, влияние от его изменения на x постоянно во всех периодах и у всех экономических объектов, отличаться может лишь среднее значение.

f_i – индивидуальные особенности, которые не зависят от времени.

γ_t – специфические особенности, которые наблюдаются не во всех периодах, но у всех объектов.

Последние две величины являются ненаблюдаемыми.

Ошибки – это случайные величины, которые независимы и одинаково распределены (математическое ожидание = 0, а дисперсия = σ_ε^2).

Временные эффекты учитываются с помощью фиктивных переменных в коротких панельных данных (с небольшим количеством периодов).

Пусть имеется рассматриваются два периода ¹¹ ($T=2$). Фиктивная переменная

$$Z^d = (Z_1^d Z_2^d), \text{ где } (Z_k^d)_{it} = \begin{cases} 1, & k = t \\ 0, & k \neq t \end{cases}.$$

Следовательно, модель можно записать в таком виде:

¹¹Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 164.

$$x_{it} = Z_{it} * a + Z^d * \gamma_t + f_i + \varepsilon_{it} \quad (2.24)$$

Существуют такие факторы, которые не изменяются по периодам. В таком случае, можно рассмотреть два типа моделей, в зависимости от сделанных предположений о коэффициенте f_i :

- Модели с фиксированными эффектами (f_i - N-постоянных неизвестных параметров),
- Модели со случайными эффектами. (f_i - случайны, не коррелированы с ошибками).

2.3.2 Модели с фиксированными и случайными эффектами

Модель с фиксированными эффектами

Данная модель является моделью линейной регрессии, её свободные переменные меняются по индивидуальным объектам i .

$$x_{it} = Z_{it} * a + f_i + \varepsilon_{it} \quad (2.25)$$

Где выполняются предположения:

Z_{it} – не зависимы от ошибок (ε_{it}).

Ошибки – случайные величины, которые являются независимыми и одинаково распределенными, где математическое ожидание равно нулю, а дисперсия равна δ_ε^2 .

Далее, модель можно записать с помощью фиктивных переменных для каждого экономического объекта.

$$x_{it} = Z_{it} * a + \sum_{j=1}^N z_{ij}^d * f_j + \varepsilon_{it} \quad (2.26)$$

Где: $Z_{ij}^d = \begin{cases} 1, i = j \\ 0, i \neq j \end{cases}$

Можно заметить, что в модели имеется N – штук фиктивных переменных. Объединив все фиктивные переменные, можно построить матрицу - Z^d , размерность (NT*N).

$$Z_d = I_N \mathbf{1}_T \otimes,$$

Где:

1_T – этот вектор-столбец, который состоит из единиц.

1_N – это единичная матрица (N)

\otimes - произведение Кронекера.

Уравнение, записанное в матричной форме, имеет следующий вид:

$$x = Z * a + Z^d * f + \varepsilon \quad (2.27)$$

Где:

$x = (x_1, x_2, \dots, x_N)'$ - зависимая переменная (вектор-столбец размерности NT)

$Z = (Z_1', Z_2', \dots, Z_N')$ - наблюдения за зависимой переменной (матрица размерности NT*n)

$\varepsilon = (\varepsilon_1', \varepsilon_2', \dots, \varepsilon_N')$ - ошибки (вектор-столбец размерности NT)

$f = (f_1', f_2', \dots, f_N')$ - параметры регрессии (вектор-столбец размерности N).

Оценки, полученные в результате применения МНК, называются МНК-оценки с фиктивными переменными. Если выполнялись предположения, записанные ранее, то такие оценки – несмещенные и эффективные. Оценивать с помощью МНК достаточно сложно с точки зрения вычислений, существуют более простые методы.

При помощи центрированных переменных возможно не использовать константу и, соответственно, значительно уменьшить число неизвестных параметров. Значит, оценка МНК для вектора a получится, при условии построения регрессии в отклонениях от индивидуальных средних значений. По факту – происходят преобразования, в следствие которых исключаются специфические эффекты f_i .

Далее, рассчитываются средние величины:

$$\bar{x}_i = \bar{Z}_i * a + \bar{f}_i + \bar{\varepsilon}_i \quad (2.28)$$

Где:

$$\bar{x}_i = \frac{1}{T} * \sum_{i=1}^T x_i, \quad \bar{Z}_i = \frac{1}{T} * \sum_{i=1}^T Z_i, \quad \bar{\varepsilon}_i = \frac{1}{T} * \sum_{i=1}^T \varepsilon_i$$

Если поэлементно вычесть данное выражение $\bar{x}_i = \bar{Z}_i * a + \bar{f}_i + \bar{\varepsilon}_i$ из $x_{it} = Z_{it} * a + f_i + \varepsilon_{it}$, то можно получить уравнение регрессии, в отклонениях от индивидуальных средних значений.

$$x_{it} - \bar{x}_i = (Z_{it} - \bar{Z}_i) * a + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

Последнее уравнение было получено после преобразований в наблюдениях из уравнения регрессии в матричной форме с помощью матрицы W, которая имеет размерность (NT*NT).

$$W_x = W * Z_a + W * Z^d * f + W_\varepsilon$$

Нетрудно показать, что данное соотношение можно записать в виде:

$$W_x = W * Z_a + W_\varepsilon \quad (2.29)$$

МНК-оценка вектора параметров а имеет вид:

$$a = ((WZ)'WZ)^{-1}(WZ)^{-1}Z'W_x = (Z'WZ)^{-1}Z'W_x$$

Такая оценка в точности совпадает с МНК-оценкой модели с фиктивными переменными. Для оценок индивидуальных особенностей часто используется:

$$f_i = \bar{x}_i - \bar{Z}_i * a_{EF}, i = 1, \dots, N.$$

Если Т стремится к бесконечности, то данные оценки состоятельные и несмещенные, а если Т постоянно, то \bar{x}_i и \bar{Z}_i при $(N \rightarrow \infty)$ не сходятся, следовательно, оценка f_i является не состоятельной. Можно проверить и доказать, что¹² если ошибки ε имеют многомерное нормальное распределение, то оценки с фиксированными эффектами являются асимптотически нормальными, следовательно, можно пользоваться стандартными процедурами для проверки гипотез относительно параметров регрессионной модели с фиксированными эффектами (t-тесты, F-тесты, тест Вальда).

Проведём тест на наличие специфических (индивидуальных) эффектов с помощью наложения ограничений на коэффициенты модели:

$$H_0 : f_1 = f_2 = \dots = f_N = 0^{13}.$$

Если ошибки независимые случайные величины, распределенные по нормальному закону с нулевым средним и дисперсией $= \sigma_\varepsilon^2$, если нулевая гипотеза верна, то

¹² Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 170.

¹³ Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 173.

$$F = \frac{(e_1' * e_1 - e' * e) / (N - 1)}{e' * e / (N * T - N - k)} \sim F_{N-1, NT-N-k}$$

Где:

e – это вектор остатков модели с фиксированными эффектами.

e_1 – это вектор остатков модели с ограничением: $x_{it} = Z_{it} * a + \varepsilon_{it}$

Если верна нулевая гипотеза, то нужно оценивать объединенную модель:

$$x_{it} = Z_{it} * a + \varepsilon_{it} \quad (2.30)$$

Модель со случайными эффектами

В моделях такого вида делается предположение, что индивидуальные различия – случайны, таким образом, регрессионное уравнение имеет следующий вид:

$$x_{it} = Z_{it} * a + u_{it} \quad (2.31)$$

u_{it} – это сумма ε_{it} и f_{it} .

Делаются предположения¹⁴ о том, что:

- Z_{js} – не зависят от ε_{it} при всех i, j, s, t ,
- ошибки — случайные величины, независимы и одинаково распределенные с нулевым математическим ожиданием и дисперсией $= \sigma_{\varepsilon}^2$,
- Z_{js} – не зависят от f_i при всех i, j, s ,
- ошибки ε_{it} и f_i не зависят друг от друга при всех i, j, s, t .

При сделанных предположениях оценки традиционного МНК являются несмещенными и состоятельными, но неэффективными.

Ввиду того, что дисперсия u_{it} равна дисперсии суммы f_i и ε_{it} ,

$$D(f_i + \varepsilon_{it}) = \sigma_f^2 + \sigma_{\varepsilon}^2$$

Использование для оценки традиционный МНК не позволяет получить наилучшие линейные несмещённые оценки параметров регрессионного уравнения. Следовательно,

¹⁴ Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 174.

для того, чтобы получить эффективные оценки, нужно использовать обобщённый метод наименьших квадратов. Рассмотрим применение данного метода. В матричном и скалярном виде преобразования можно выразить следующим образом:

В матричном виде:

$$Dx = DZa + Du \quad (2.32)$$

В скалярном виде:

$$x_{it} - \Delta * \bar{x}_i = (Z_{it} - \Delta * \bar{Z}_i)a + u_{it} - \Delta * \bar{u}_i \quad (2.33)$$

Где $\Delta \in [0,1]$

Таким образом, оценка, полученная с помощью обобщенного МНК для вектора a выглядит следующим образом:

$$a_{GLS} = \left[\sum_{i=1}^N \sum_{t=1}^T (Z_{it} - \Delta * \bar{Z}_i)' (Z_{it} - \Delta * \bar{Z}_i) \right]^{-1} \left[\sum_{i=1}^N \sum_{t=1}^T (Z_{it} - \Delta * \bar{Z}_i)' (x_{it} - \Delta * \bar{x}_i) \right]$$

Оценка такого вида называется оценкой со случайными эффектами.

Если $\Delta = 0$, то оценки Обобщенного МНК и Традиционного совпадают. Если $\Delta = 1$, то оценка модели со случайным эффектом совпадает с оценкой модели с фиксированными эффектами. На практике, для того, чтобы применить Обобщенный МНК для модели, которая содержит случайные эффекты, требуется знать величину Δ . Часто, эта величина не известна. Для того, чтобы получить оценку параметров модели, нужно оценивать дисперсии ошибок σ_f^2 и σ_ε^2 . Требуется оценить внутригрупповую регрессию, вычислив для неё остатки, далее – оценить дисперсию δ_ε^2 .

$$\hat{s}_\varepsilon^2 = \frac{1}{N * T - N - n} RSS_w$$

$$\hat{s}_f^2 + \frac{\hat{s}_\varepsilon^2}{T} = \frac{RSS_B}{N - n}$$

Далее, для дисперсии σ_f^2 можно выразить оценку:

$$\hat{s}_f^2 = \frac{RSS_B}{N - n} - \frac{\hat{s}_\varepsilon^2}{T} \quad (2.34)$$

С помощью последнего выражения можно вычислить величину Δ и применить Обобщенный МНК.

2.3.3 Выбор наилучшей модели

Как известно, для обыкновенных моделей одним из показателей качества является коэффициент детерминации (обычный или исправленный), который трактуется как отношение объясненной дисперсии в общей. Но в моделях, которые содержат панельные данные, вариация во внутригрупповых и межгрупповых моделях оценивается по-разному.

Сумма внутригрупповой и межгрупповой дисперсии даёт полную дисперсию.

$$\frac{1}{N * T} \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x})^2 = \frac{1}{N * T} \sum_{i=1}^N \sum_{t=1}^T (x_{it} - \bar{x}_i)^2 + \frac{1}{N} \sum_{i=1}^N (\bar{x}_i - \bar{x})^2$$

Для оценки качества модели с фиксированными эффектами R^2 нужно выбрать таким образом, чтобы внутригрупповая дисперсия объяснялась наиболее полно.

$$R_w^2 = r^2(x_{it} - \bar{x}_i, x_{it}^e - \bar{x}_i^e),$$

Где: r^2 – это квадрат парной корреляции.

$$x_{it}^e - \bar{x}_i^e = (Z_{it} - \bar{Z}_i) a_{FE}$$

Коэффициент аппроксимации данных для межгрупповой регрессии рассчитывается следующим образом:

$$R_B^2 = r^2(x_{it}, x_{it}^e) \quad (2.35)$$

Где: $x_{it}^e = Z_{it} * a_B$

Для того чтобы оценить точность обычной модели, R^2 вычисляется следующим образом:

$$R_{overall}^2 = r^2(x_{it}, x_{it}^e)$$

Где: $x_{it}^e = Z_{it} * a_{OLS}$.

Все три варианта коэффициента детерминации можно рассчитать для любой регрессионной модели, оцененной с помощью любого метода. R^2 , оценённый для модели со случайными эффектами, всегда строго меньше, чем R^2 , оцененный для модели с фиксированными эффектами. Соответственно, не следует использовать R^2 , выбирая между способами оценивания модели. Но, R^2 целесообразно использовать для сравнения нескольких моделей, которые отличаются набором регрессоров, но оцененные одинаковыми методами.

Выбирая способ оценивания между общей моделью, моделью с фиксированными или случайными эффектами, возникает вопрос, какую же модель выбрать.

Содержательно, разница между моделями заключается в том, что обычная регрессия – у экономических объектов нет индивидуальных (специфических) различий. Модель с фиксированными эффектами предполагает, что каждый экономический объект – особый и не рассматривается как случайная выборка. Но если экономические единицы попали в выборку случайно, выбирая из большой совокупности, то следует использовать модель со случайными эффектами.

Кроме общих представлений о модели, нередко используются статистические тесты, которые позволяют ответить на вопрос о выборе модели. Нулевой гипотезой является более частная модель, а альтернативной – более общая¹⁵.

1. Общая регрессия это частный случай модели с фиксированными эффектами, если в модели с фиксированными эффектами $f_1 = f_2 = \dots = f_N = 0$.
2. Обычная регрессия – частный случай модели со случайными эффектами, когда в модели со случайными эффектами $\sigma_f^2 = 0$.
3. Модель со случайными эффектами – частный случай модели с фиксированными эффектами, если в модели с фиксированными эффектами f_i не коррелирует с регрессорами Z_{it} .

Выбирая между моделью с фиксированными эффектами и обычной, используется F-тест, в котором накладывается условие $f_1 = f_2 = \dots = f_N = 0$. Используется для модели с

фиктивными переменными: $x_{it} = Z_{it} a + \sum_{j=1}^N z_{it}^d * f_i + \varepsilon_{it}$.

Выбирая между обычной моделью и моделью со случайными эффектами, следует проводить тест Брюшта-Пагана¹⁶ для $\sigma_f^2 = 0$.

$$LM = \frac{N * T}{2 * (T - 1)} \left(\frac{e' * Z^d * Z^{d'} * e}{e' * e} - 1 \right)^2 \sim \chi^2$$

При $\sigma_f^2 = 0$, где e – это остатки обычной регрессии, а Z^d – матрица фиктивных переменных.

¹⁵ Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 185.

¹⁶ Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - С 186.

Если стоит выбор между моделью со случайными и фиксированными эффектами, следует провести тест Хаусмана, сравнив оценки регрессий внутригрупповой и со случайными эффектами. Если данные оценки схожи, то нужно выбирать модель со случайными эффектами, потому что она содержит меньшее число параметров регрессии, следовательно, она проще. Нулевой гипотезой является модель со случайными эффектами.

$$H = (a_{GLS} - a_w)' = \text{var}(a_{GLS} - a_w)^{-1} (a_{GLS} - a_w) \sim \chi^2_{H_0}$$

2.4 Построение и анализ моделей панельных данных

В качестве примера будет рассмотрена модель зависимости логарифма валового регионального продукта (grp) от логарифма прямых инвестиций (Invest), среднего уровня заработной платы (income) и среднего уровня занятости (employed). Рассматриваются восемь федеральных округов Российской Федерации, за период с 2000 по 2014 год.

$$l_{GRP} = \text{const} + a_1 * l_{Invest} + a_2 * income + a_3 * employed + \varepsilon$$

Оценив данную модель с помощью обычного МНК, получаются следующие коэффициенты, тестовые статистики и уровни значимости.

	Коэффициент	Ст.ошибка	t-статистика	Р-значение
const	3,6058	0,350477	10,2883	<0,00001 ***
employed	5,45359e-05	3,64051e-06	14,9803	<0,00001 ***
income	1,14543e-05	3,4013e-06	3,3676	0,00103 ***
l_Invest	0,785479	0,0308711	25,4438	<0,00001 ***

Проведём тест Уайта на наличие гетероскедастичности.

Нулевая гипотеза: гетероскедастичность отсутствует

Тестовая статистика: LM = 16,9277

р-значение = P(Chi-квадрат(9) > 16,9277) = 0,0498606

Соответственно, нулевая гипотеза отвергается и в модели присутствует гетероскедастичность.

Проведём тест на избыточные переменные.

Нулевая гипотеза: параметры регрессии нулевые для следующих переменных: employed, income

Тестовая статистика: $F(2, 116) = 124,333$

$p\text{-значение} = P(F(2, 116) > 124,333) = 1,40813\text{e-}029$

Т.к. $p\text{-значение}$ равно нулю, то нулевую гипотезу можно отвергнуть. Соответственно, параметры регрессии *employed* и *income* – ненулевые.

Далее, проведём оценку данной модели с точки зрения фиксированных эффектов.

Такой способ оценивания предполагает наличие у федеральных округов индивидуальных характеристик, которые не меняются во времени. Такие характеристики бывает невозможно измерить или даже наблюдать. Если такие характеристики являются ненаблюдаемыми, то их можно включить в уравнение в виде объясняющих переменных. Для оценивания добавим переменные *dummy* (фиктивные).

$$\begin{cases} d_{ij} = 1, \text{при } i = j \\ d_{ij} = 0, \text{при } i \neq j \end{cases}$$

Оцениваемая модель принимает вид:

$$l_{GRP} = const + \sum_{j=1}^N a_j * du_{ij} + a_1 * l_{Invest} + a_2 * income + a_3 * employed + \varepsilon$$

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	4,7262	0,287517	16,4380	<0,00001	***
du_1	-1,76786	0,82035	-2,1550	0,03336	**
du_2	-0,279629	0,187739	-1,4895	0,13926	
du_3	-0,417826	0,154616	-2,7024	0,00799	***
du_4	-0,119814	0,0353245	-3,3918	0,00097	***
du_5	-1,39033	0,601634	-2,3109	0,02272	**
du_6	-0,0728025	0,145463	-0,5005	0,61774	
du_7	-0,437354	0,307515	-1,4222	0,15782	
l_Invest	0,652606	0,0274004	23,8174	<0,00001	***
employed	0,000182936	5,48347e-05	3,3361	0,00116	***
income	2,05368e-05	2,64505e-06	7,7642	<0,00001	***

В результате оценивания из-за совершенной коллинеарности была пропущена переменная *du_8*. Аналогично можно оценить модель с фиксированными эффектами, встроенную в Gretl.(приложение 8)

Нулевая гипотеза теста на различие констант в группах отвергается, соответственно, группы не имеют общие константы.

Проведя тест на нормальность остатков, можно сделать вывод, что остатки не имеют нормального распределения. Вероятность ошибиться, отвергая нулевую гипотезу (нулевая гипотеза – распределение является нормальным), составляет 0,0397, это меньше пяти

процентного уровня значимости, следовательно, нулевая гипотеза отвергается. На рисунке 2.2 представлен график распределения остатков модели.

Далее будет оценена та же зависимость, но только с помощью модели со случайными эффектами. Результаты расчетов представлены в приложении 9. Все переменные являются значимыми на пяти процентном уровне значимости.

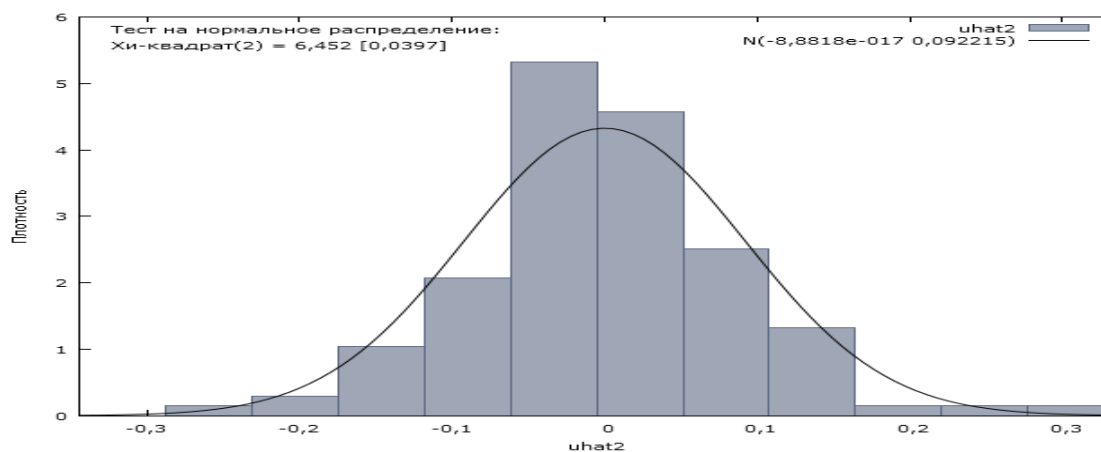


Рисунок 2.2 Тест на нормальное распределение

В модели рассчитываются внутригрупповые и межгрупповые дисперсии. Обобщенная оценка наименьших квадратов в такой модели учитывает и изменчивость межгрупповую, и внутригрупповую. Данная оценка – это средневзвешенная оценка межгрупповой и внутригрупповой оценок. Обе оценки являются состоятельными, соответственно, обобщенная оценка – также состоятельна.

Далее будет проведено сравнение трёх полученных моделей и выбор наиболее адекватной из них.

1. Проводится *тест Вальда* (в модели с фиксированными эффектами) на равенство нулю всех индивидуальных эффектов.

Нулевая гипотеза: параметры регрессии нулевые
 $du_1, du_2, du_3, du_4, du_5, du_6, du_7$

Тестовая статистика: $F(7,109) = 30.553$, P-значение = $5.177e-0.23$

Полученная статистика отвергает нулевую гипотезу по тесту Вальда. Соответственно, для описания имеющихся данных обычная регрессионная модель хуже, чем модель с фиксированными эффектами.

2. Проводится *тест Бриша-Пэгана* на наличие случайного индивидуального эффекта.

Нулевая гипотеза: Дисперсия специфических для наблюдений ошибок равна нулю
Асимптотическая тестовая статистика Хи-квадрат(1) = 252.601, Р-значение = 7.035e-0.57

Р-значение примерно равно нулю, соответственно, нулевая гипотеза отвергается. Модель со случайными эффектами лучше описывает данные, чем обычная регрессионная модель.

3. Проведение *теста Хаусмана* для сравнения моделей со случайным и фиксированным эффектами.

Значение тестовой статистики равно 8.53 (Р-значение равно 0.04). Соответственно, отвергается нулевая гипотеза о том, что корреляция между ошибками и регрессорами отсутствует (равна нулю). Следовательно, выбирая между моделями со случайными и фиксированными эффектами, выбирается модель с фиксированными эффектами.

Проведя сопоставление моделей можно сделать вывод, что в рассмотренном случае наилучшим образом подходит модель с фиксированными эффектами. Этот вывод ожидаемый, т.к. для изучения выбирались федеральные округа, состав которых не изменялся в течение рассматриваемых лет.

2.5 Выбор наиболее предпочтительного округа

Большая часть жителей небольших городов (особенно молодёжь) стоят перед выбором – сменить или нет своё постоянное место жительства, уехать «на заработки» в другие регионы. Соответственно, перед ними встаёт вопрос о выборе наиболее привлекательного региона для жизни, работы и создания семьи. По каждому федеральному округу рассматривается шесть характеристик: инвестиции в основной капитал, количество зарегистрированных больных на 1 000 человек населения, уровень зарегистрированных преступлений на 100 000 человек населения, уровень миграции, среднедушевые денежные доходы населения, уровень занятости.

Предположим, что человек стоит перед выбором наиболее привлекательного региона (Федерального Округа). Для него важными характеристиками являются – здоровье – как часто болеет население в данном регионе, инвестиции – на сколько регион поддерживается, и финансируются ли предприятия данного региона, уровень занятости – сможет ли данный человек устроиться на работу, доходы – на какую в среднем заработную плату он сможет рассчитывать, преступность – безопасно ли в данном

регионе, какой в ней уровень преступности и уровень миграции – больше въезжают в данный регион или выезжают.

Целью данного изучения является выбор наиболее привлекательного региона с помощью анализа характеристик Федеральных Округов, а также анализа по методологии АСПИД (АСПИД – «Анализ и Синтез Показателей при Информационном Дефиците»), изложенной в работе Н.В.Хованова¹⁷. Преимущества данного метода заключаются в том, что возможно учитывать всю нечисловую, неточную и неполную информацию, которой располагает исследователь, а также результаты, полученные в процессе обработки имеющейся информации, наглядны.

В ходе исследования будет проанализировано четыре модели. *Первая модель* – без интервальных и порядковых предположений о весовых коэффициентах и сводных показателях предпочтительности. *Вторая модель* (Синтез) – с учётом экспертных оценок о весовых коэффициентах. *Третья модель* (Анализ) – с учётом порядковой информации об агрегированных показателях предпочтительности. *Четвертая модель* - Анализ и Синтез, включает в себя порядковую информацию о весовых коэффициентах и об агрегированных показателях предпочтительности. Рассмотрим показатели, на основе которых проводилось исследование.

1. Инвестиции (объём прямого инвестирования в регион млн.руб).
2. Уровень преступности (Рассматривается численность официально зарегистрированных преступлений на 100 000 человек населения).
3. Среднедушевой доход (тыс.руб).
4. Уровень здравоохранения (Рассматривается количество зарегистрированных заболеваний у пациентов с диагнозом, который установлен впервые в их жизни. Рассчитывается на 1000 человек населения).
5. Уровень миграции (чем ниже показатель, тем лучше).
6. Уровень занятости (чем выше, тем лучше).

В следующей таблице 2.2 представлены данные по всем переменным (признакам) и Федеральным округам (альтернативам) на основе Региональной статистики РФ. Использовались данные за 2014 год.

¹⁷ Хованов Н.В, Анализ и синтез показателей при информационном дефиците/ Н.В. Хованов, СПбГУ, 1996г – 196 с.

Таблица 2.2 Оценка признаков объектов

Признак Альтернатива	Инвестиции	Занятость	Доход	Здоровье	Ур. миграции	Преступность
Центральный	3435974	19008,3	34970	715	56	1499
Северо-Западный	1357860	6750,2	28572	854	41	1474
Южный	1277238	6161,2	24328	712	34	1303
Северо-Кавказский	516921	3464,1	20692	670	-21	724
Приволжский	2355973	14114,8	24020	872	-2	1352
Уральский	2322596	6037,1	30494	817	7	1764
Сибирский	1440980	9010,1	21490	861	-4	2038
Дальневосточный	820142	3267,5	31974	805	-40	2092

В таблице 2.3 рассматриваются граничные значения каждого признака (их Min и Max значения), рассчитаны средние значения и стандартные отклонения.

Таблица 2.3 Статистические значения признаков и нормировка

N	Название признака	Min	Max	Среднее	Ст.отклон.	Inc / Dec
1	Инвестиции	516921	3435974	1690960,5	892159,1263	Inc
2	Занятость	3267,5	19008,3	8476,6625	5117,9	Inc
3	Доход	20692	34970	27067,5	4861,5236	Inc
4	Здоровье	670	872	788,25	73,2287	Dec
5	Уровень миграции	-40	56	8,875	30,5223	Dec
6	Преступность	724	2092	1530,75	413,1594	Dec

«Inc» и «Dec» определяют то, каким образом оценивается наш признак. «Inc» означает, что чем больше значение признака, тем оно более привлекательно. «Dec» – наоборот – чем ниже значение признака – тем лучше. Соответственно, чем выше уровень инвестирования, занятости и среднего уровня заработка в Федеральном округе, тем данный округ более привлекателен. И чем ниже уровень заболеваемости, миграции и преступности, тем лучше данный округ.

Рассмотрим таблицу 2.4. Единица означает, что данный признак имеет наилучшее значение в данном округе, а ноль – наихудшее. Таким образом, можно сделать вывод о том, что *Центральный ФО* является наиболее привлекательным по показателям – инвестиции, занятость и доход. Но по уровню миграции данный регион является наихудшим (уровень миграции в ЦФО является максимальным). *Северо-Кавказский ФО* по уровню инвестиций и доходу является наименее привлекательным, но по уровню здравоохранения и уровню преступности округ является благоприятным (т.к. наименьшее число заболевших и уровень преступности меньше остальных регионов).

Таблица 2.4 Нормированные показатели и их описательная статистика

Альтернатива	Инвестиции	Занятость	Доход	Здоровье	Ур. миграции	Преступность
Центральный	1,0000	1,0000	1,0000	0,7772	0,0000	0,4335
Северо-Западный	0,2881	0,2213	0,5519	0,0891	0,1563	0,4518
Южный	0,2605	0,1838	0,2547	0,7921	0,2292	0,5768
Северо-Кавказский	0,0000	0,0125	0,0000	1,0000	0,8021	1,0000
Приволжский	0,6300	0,6891	0,2331	0,0000	0,6042	0,5409
Уральский	0,6186	0,1760	0,6865	0,2723	0,5104	0,2398
Сибирский	0,3166	0,3648	0,0559	0,0545	0,6250	0,0395
Дальневосточный	0,1039	0,0000	0,7902	0,3317	1,0000	0,0000

Модель 1.

Для исходной модели, которая будет построена изначально – начальные интервалы весовых коэффициентов – заданы автоматически, т.е. вес каждого признака между 0 и 1. Также не будут заданы предпочтительные регионы или признаки.

Соответственно, исходя из сделанных предположений можно сформулировать следующие результаты. В таблице 2.5 представлены расчётные данные по весам. Среднее значение каждого веса равно 0,1667 возможные отклонения от расчётного значения задаются с помощью стандартного отклонения ($\pm 0,1450$).

Таблица 2.5 Весовые коэффициенты оценок

Вес признака	Min	Max	Среднее	Станд. отклон.
w (Инвестиции)	0,0000	1,0000	0,1667	0,1450
w (Занятость)	0,0000	1,0000	0,1667	0,1450
w (Доход)	0,0000	1,0000	0,1667	0,1450
w (Уровень урбанизации)	0,0000	1,0000	0,1667	0,1450
w (Уровень миграции)	0,0000	1,0000	0,1667	0,1450
w (Образование)	0,0000	1,0000	0,1667	0,1450

В таблице 2.6 представлены результаты расчётов при заданных начальных условиях. Среднее значение, рассчитанное в таблице 2.6, и есть искомый агрегированный показатель предпочтительности. Стандартное отклонение показывает возможный разброс данного показателя влево и вправо.

Таблица 2.6 Статистика агрегированного показателя предпочтительности

Альтернатива	Min	Max	Среднее	Ст.отклон.	Место
Q (Центральный Ф.О.)	0,0000	1,0000	0,7018	0,1453	1
Q (Северо-Западный Ф.О.)	0,0891	0,5519	0,2931	0,0631	7
Q (Южный Ф.О.)	0,1838	0,7921	0,3828	0,0870	5
Q (Северо-Кавказский Ф.О.)	0,0000	1,0000	0,4691	0,1827	2
Q (Приволжский Ф.О.)	0,0000	0,6891	0,4496	0,0968	3
Q (Уральский Ф.О.)	0,1760	0,6865	0,4173	0,0766	4
Q (Сибирский Ф.О.)	0,0395	0,6250	0,2427	0,0838	8
Q (Дальневосточный Ф.О.)	0,0000	1,0000	0,3710	0,1523	6

В таблице 2.7 рассчитаны корреляции сводного показателя предпочтения. Можно заметить, что при увеличении предпочтительности Центрального Ф.О. также увеличивается предпочтительность Северо-Западного Ф.О, Южного Ф.О и Уральского Ф.О, но предпочтительность снижается в Северо-Кавказском Ф.О, Приволжском, Сибирском и Дальневосточном Ф.О. Таблица 2.7 является симметричной. наиболее сильная зависимость наблюдается между Центральным и Северо-Кавказским Ф.О (≈ -1), также между Сибирским и Приволжским ($\approx +0,9$). Практически не коррелируют Северо-Западный и Дальневосточный федеральные округа (коэффициент корреляции равен минус 0,05).

Таблица 2.7 Корреляция сводного показателя предпочтительности

QCORR(i,j)	Q (Центральный)	Q (Северо- Западный)	Q (Южный)	Q (Северо- Кавказский)	Q (Приволжский)	Q (Уральский)	Q (Сибирский)	Q (Дальневосточн.)
Q (Центральный)	1,00	0,58	0,26	-0,99	-0,36	0,02	-0,56	-0,80
Q (Северо-Западный)	0,58	1,00	0,87	-0,62	0,15	0,55	0,21	-0,05
Q (Южный)	0,26	0,87	1,00	-0,32	0,63	0,88	0,62	0,26
Q (Сев.-Кавказский.)	-0,99	-0,62	-0,32	1,00	0,32	-0,07	0,51	0,77
Q (Приволжский)	-0,36	0,15	0,63	0,32	1,00	0,88	0,90	0,58
Q (Уральский)	0,02	0,55	0,88	-0,07	0,88	1,00	0,81	0,44
Q (Сибирский)	-0,56	0,21	0,62	0,50	0,90	0,81	1,00	0,86
Q (Дальневосточный)	-0,80	-0,05	0,26	0,77	0,58	0,44	0,86	1,00

Модель 2. (Синтез)

Далее, введем наши предположения о весовых коэффициентах. Пусть для нас важнее здоровье, чем миграция. $w(\text{Здоровье}) > w(\text{Миграция})$. Соответственно, произойдут следующие изменения. На рисунке 2.3 отражены введенные соотношения: вес здоровья выше веса миграции.

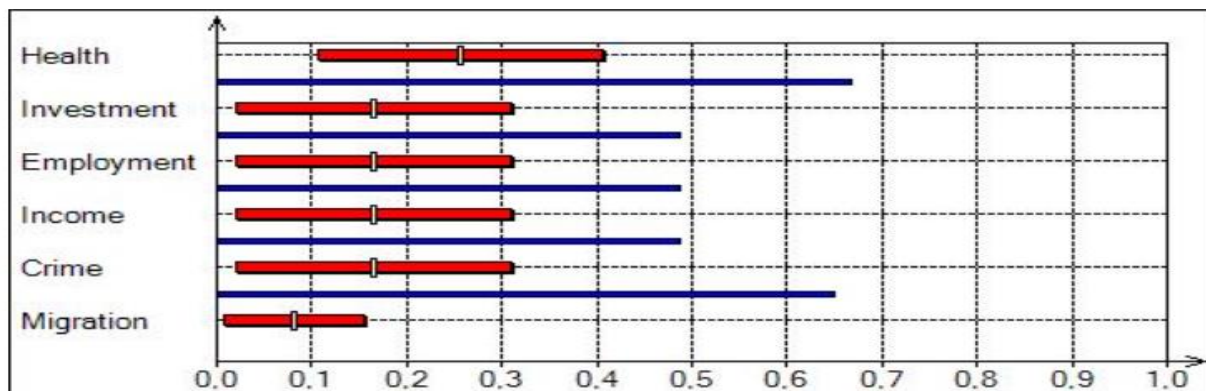


Рисунок 2.3 Весовые коэффициенты оценок

Исходя из данных изменений, формируется искомый показатель предпочтительности.

Модель 3 (Анализ)

Далее, добавим к исходной модели новые предположения:

Первое предположение о том, что Дальневосточный Федеральный Округ предпочтительнее, чем Сибирский ввиду его выгодного приморского положения, а также близости к странам Азиатско-Тихоокеанского региона, которые отличаются интенсивным экономическим ростом. Важным преимуществом Дальневосточного федерального округа является наличие крупных городов и портов, транспортных путей в южной части округа.
 $Q(\text{Дальневосточный Ф.О.}) > Q(\text{Сибирский Ф.О.})$

Второе предположение – Южный Федеральный Округ предпочтительнее Северо-Кавказского, ввиду близости к Черному морю и большого количества рекреационных зон, высокой социально-экономической привлекательности, сформированной спортивной и транспортной инфраструктуры (Сочинский олимпийский комплекс), и т.д.

$$Q(\text{Южный Ф.О.}) > Q(\text{Северо-Кавказский Ф.О.})$$

Таким образом, распределение весовых коэффициентов представлены на следующем рисунке 2.4, получаем, что вес дохода с вероятностью 0,68 превосходит все остальные веса.

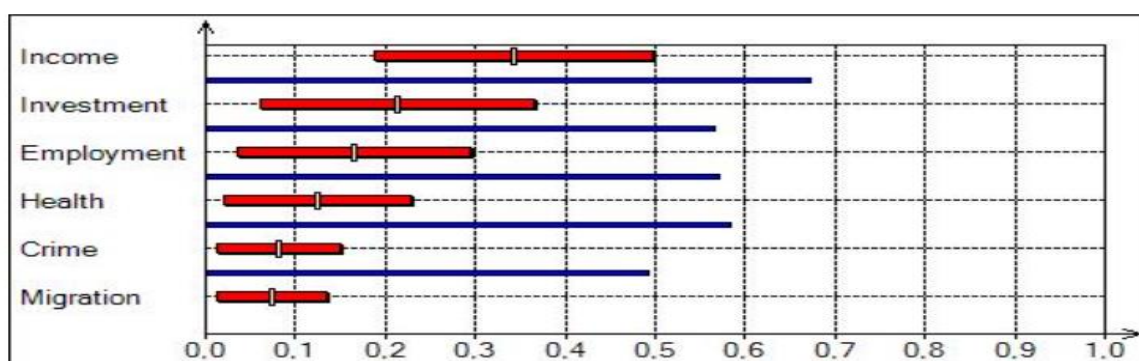


Рисунок 2.4 Весовые коэффициенты оценок

Модель 4 (Анализ и Синтез)

Изменим исходную модель 1, дополнив её уже известными предположениями:

- 1) $w(\text{Здоровье}) > w(\text{Миграция})$
- 2) $Q(\text{Дальневосточный Ф.О.}) > Q(\text{Сибирский Ф.О.})$
- 3) $Q(\text{Южный Ф.О.}) > Q(\text{Северо-Кавказский Ф.О.})$.

Таким образом, будет проводиться процедура Анализа и Синтеза. Получим следующие распределения весовых коэффициентов.

При всех введенных новых условиях получаются следующие результаты и распределения весовых коэффициентов.

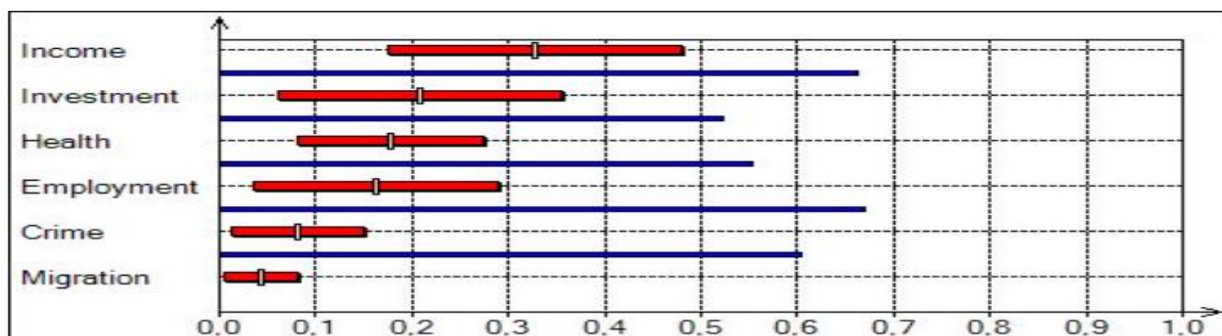


Рисунок 2.5 Весовые коэффициенты оценок

Соответственно, построив четыре модели, можно сформулировать выводы по каждой из моделей, становится возможным провести их сравнение, выделение общих закономерностей.

Выводы по четырём моделям

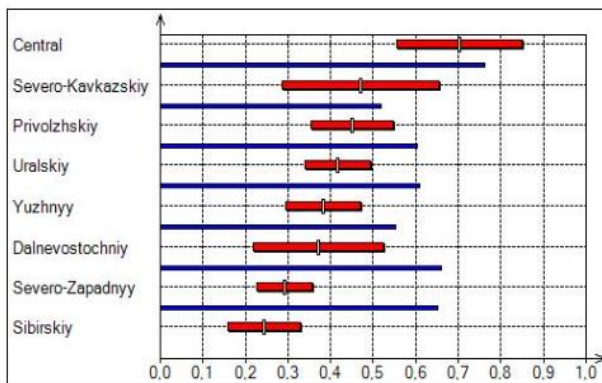
В модели 1 наиболее благоприятным является Центральный Федеральный округ. С вероятностью 0,76 он доминирует над всеми остальными округами. Далее, с вероятностью 0,52 Северо-Кавказский ФО лучше, чем остальные (за исключением ЦФО). Приволжский ФО с вероятностью 0,6 превосходит оставшиеся округа. Северо-Западный ФО с вероятностью 0,66 предпочтительнее наименее привлекательного – Сибирского Ф.О. Данный рисунок можно рассматривать также как сортировку, в который сверху вниз представлены Федеральные округа по степени предпочтительности (чем ниже, тем менее привлекателен).

В модели 2 при сопоставлении полученных результатов в первом и втором случае, наиболее и наименее предпочтительные регионы не изменились. Поменялись лишь 3,4,5 места. Приволжский Ф.О опустился на одно место, заняв 3 место, Южный Ф.О поднялся на две строчки, заняв 4 место, а Уральский Ф.О опустился на одну строку.

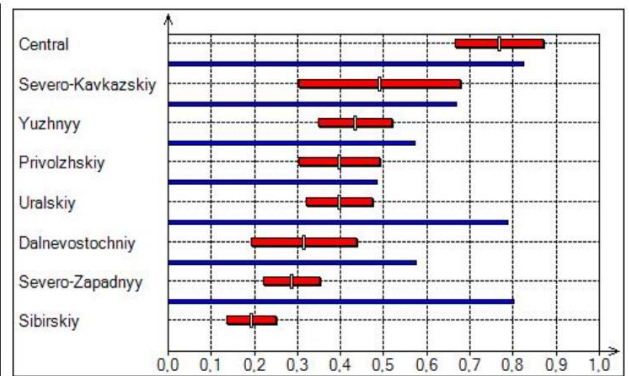
В модели 3 полученные результаты свидетельствуют о том, что Центральный Ф.О. с вероятностью 100% предпочтительнее всех остальных Округов. Северо-Кавказский Ф.О. опустился со второго места на 7, Северо-Западный Ф.О. поднялся с 7 места на 5, Уральский Ф.О. занял второе место.

В модели 4 Центральный Федеральный Округ является лидером. Со 100% вероятностью можно утверждать, что при имеющихся предположениях и прочих равных условиях Центральный Ф.О. предпочтительнее всех остальных округов. Второе место занимает Уральский Федеральный Округ. Если сравнивать исходную модель и последнюю, можно отметить, что Уральский Ф.О. поднялся с 4 места на 2. А Северо-Кавказский опустился со второго места, на предпоследнее. Северо-Западный Ф.О. поднялся на одну позицию и занял 6 место.

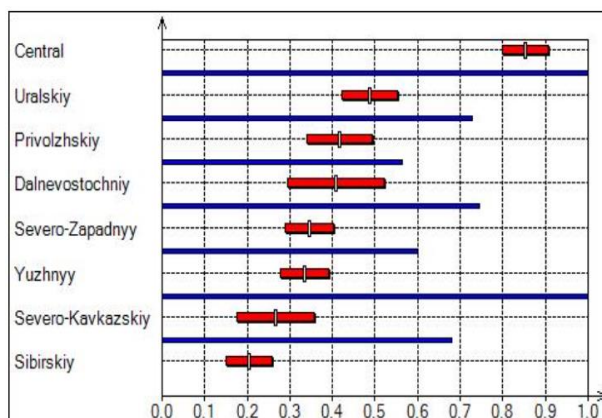
Во всех четырёх моделях наиболее предпочтительным округом является Центральный Ф.О., а наименее привлекательным – Сибирский Ф.О. Все остальные округа меняют свои позиции в процессе добавления условий о весовых коэффициентах и порядковых условиях о доминировании отдельных округов над другими. Можно заметить, что двигаясь от первой к последней модели, значительно уменьшается разброс (стандартное отклонение), наши полученные результаты становятся более точными.



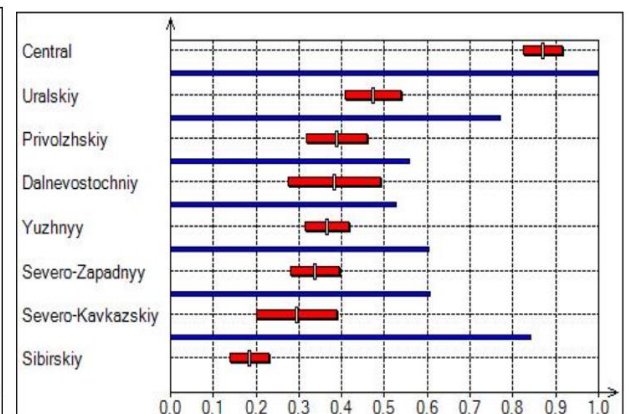
Модель 1. Без предположений



Модель 2. Предположения о w



Модель 3. Предположения о Q



Модель 4. Предположения о w и Q

Рисунок 2.6 Показатель предпочтительности

Выводы по главе II

В данной главе рассматриваются системы одновременных уравнений и анализ панельных данных.

Несмотря на то, что все федеральные округа по своей сути однородны, но даже исходя из построенных уравнений, при оценивании систем одновременных уравнений, становится видно, что развиваются федеральные округа по-разному. Можно заметить, что Центральный федеральный округ и Северо-Западный не возможно сопоставлять и сравнивать с такими федеральными округами как Южный и Уральский, т.к. разница в коэффициентах при переменных, а также при константе различаются в десятки раз.

При анализе панельных данных было построено уравнение зависимости уровня валового регионального продукта от инвестиций, среднего уровня занятости и среднедушевых денежных доходов населения. Были рассмотрены три модели – обычная регрессия, модель с фиксированными эффектами и модель со случайными эффектами. Проведя сравнение полученных моделей, был сформулирован вывод о том, что модель с фиксированными эффектами является наиболее точной.

При анализе статистических данных, полученные результаты свидетельствуют о том, что наша страна действительно имеет существенное территориальное различие по регионам в уровнях инвестирования, среднем уровне заработной платы и уровне занятости.

Также был проведён анализ Федеральных округов с точки зрения их привлекательности. В результате используемых предположений и проведенного анализа можно сформировать рейтинг привлекательности регионов. Лидером данного рейтинга является Центральный ФО, который выступает в роли наиболее привлекательного округа, за ним следуют Уральский, Приволжский, Дальневосточный, Южный, Северо-Западный и Северо-Кавказский округа, и замыкает рейтинг – Сибирский ФО.

Заключение

В работе было проанализировано социально-экономическое положение федеральных округов России, рассмотрены основные показатели, характеризующие экономическое развитие отдельных регионов. Анализируя ряд показателей, с одной стороны, можно выделить общие тенденции развития, но с другой стороны – существуют весомые различия в абсолютных значениях показателей для различных регионов.

Анализ экономических показателей федеральных округов был проведён с помощью эконометрического инструментария, а именно, систем одновременных уравнений и моделей, основанных на панельных данных, и математических методов анализа и синтеза показателей при информационном дефиците. Используемые методы позволяют оценивать достаточно сложные экономические модели, направлены на выявление взаимосвязей между отдельным явлениям и способны помочь в формировании выводов по таким экономическим процессам.

Анализ систем одновременных уравнений позволил сформулировать следующий вывод. Построенные уравнения свидетельствуют о том, что развиваются федеральные округа по-разному, несмотря на то, что все регионы однородны. Можно заметить, что Центральный федеральный округ и Северо-Западный не представляется возможным сопоставлять и сравнивать с такими округами как Южный и Уральский, т.к. разница в коэффициентах при переменных, а также константы различаются в десятки раз.

При анализе панельных данных было построено уравнение зависимости уровня валового регионального продукта от инвестиций, среднего уровня занятости и среднедушевых денежных доходов населения. Рассматривались три модели – обычная регрессия, модель с фиксированными эффектами и модель со случайными эффектами. Проведя сравнение полученных моделей, был сформулирован вывод о том, что модель с фиксированными эффектами является наиболее точной и отражает индивидуальные эффекты, которые присущи округам. Анализируя статистические данные, результаты свидетельствуют о том, что наша страна действительно имеет существенное различие по регионам в уровнях инвестирования, среднем уровне заработной платы и уровне занятости. Анализируя данную модель, была выявлена зависимость между логарифмом ВРП и логарифмом прямых инвестиций (при увеличении прямых инвестиций на один процент, уровень валового регионального продукта увеличивается на 0,65%).

На основе данных по Федеральным округам за 2014 год, полученных с помощью сайта Государственной Статистики, был построен агрегированный показатель предпочтительности округа по отдельным характеристикам, таким образом, был проведён анализ регионов с точки зрения их привлекательности. В результате используемых предположений и проведенного анализа можно сформировать рейтинг привлекательности округов. Лидером данного рейтинга является Центральный ФО, который выступает в роли наиболее привлекательного, за ним следуют Уральский, Приволжский, Дальневосточный, Южный, Северо-Западный и Северо-Кавказский округа, и замыкает рейтинг – Сибирский ФО. Таблица корреляции сводных показателей предпочтительности, построенная в процессе анализа, отражает зависимости данного показателя между округами. К примеру, корреляция сводного показателя предпочтительности между Центральным и Северо-Кавказским Ф.О. примерно равна единице, соответственно, при увеличении привлекательности Северо-Кавказского округа, снижается привлекательность Центрального, и наоборот.

Таким образом, на основе рассмотренных данных можно сделать выводы, что регионы России имеют индивидуальные особенности, развиваются неравномерно, отмечаются существенные различия между округами, ввиду большой территории страны и различных условий для проживания и развития бизнеса. Отмечаются как экономические различия – в уровне ВРП и инвестирования, так и социальные различия – дифференциация в уровне оплаты труда, численности занятых и уровне безработицы.

Список литературы

1. Анатольев С. Эконометрика для подготовленных: Курс лекций – Москва, РЭШ – 65 с.
2. Белько И. В. Эконометрика. Практикум: Учебное пособие / И. В. Белько, Е. А. Криштапович. – Минск: Изд-во Гревцова, 2011. – 224 с., ил.
3. Вербик Марно. Путеводитель по современной эконометрике. Пер. с англ. В.А.Банникова. – М: «Библиотека Солев». Научная книга, 2008. – 616с.
4. Доугерти К. Введение в эконометрику: Пер. с англ. – М.: ИНФРА-М, 1999. – XIV, 402 с.
5. Коломак Е.А. Эконометрический анализ панельных данных. — Новосибирск: НГУ, 2007
6. Кремер Н. Ш., Путко Б. А. Эконометрика: Учебник для вузов / Под ред. проф. Н. Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.
7. Лапо В.Ф. Экономика России в многополярном мире // Прикладная эконометрика. – 2013. - №30(2) – С. 26-48.
8. Магнус Я. Р., Катышев П. К., Пересецкий А. А. Эконометрика. Начальный курс: Учебник – 6-е изд., перераб. и доп. – М.: Дело, 2004. – 576 с.
9. Носко В. П. Эконометрика. Кн. 1. Ч. 1, 2: учебник / В. П. Носко. — М.: Дело РАНХиГС, 2011. — 672 с.
10. Ратникова Т.А. Введение в анализ панельных данных, ГУ- ВШЭ, 2004
11. Ратникова Т.А. Анализ панельных данных в пакете «STATA». Государственный университет Высшая школа экономики. М. 2004 – 40 с.
12. Суслов В.И., Лапо В.Ф., Талышева Л.П., Ибрагимов Н.М. Эконометрия-3: Курс лекций - Красноярск: СФУ. - 194 с
13. Хованов Н.В, Анализ и синтез показателей при информационном дефиците/ Н.В. Хованов, СПбГУ, 1996г – 196 с.
14. Виртуальная Россия: <http://virru.ru/federalnye-okruga-rossii>
15. «Группа восьми» в цифрах: Статистический сборник. (2009). М.: Росстат.
16. Регионы России. (2000–2014). Статистический сборник. М.: Росстат.
17. Российский статистический ежегодник. (2000–2014). Статистический сборник. М.: Росстат.
18. Россия и страны мира. (2014). Статистический сборник. М.: Росстат.
19. Россия и страны — члены Европейского союза. (2009). Статистический сборник. М.: Росстат.
20. The World Bank: World Development Indicators. (2015). <http://data.worldbank.org/data-catalog/worlddevelopment>
21. Wooldridge J. M. Econometric Analysis of Cross Section and Panel Data, MIT Press, 2002.
22. <http://www.finmarket.ru/main/article/3303366>. 12.04.2013
23. <http://www.iep.ru/files/persona/nosko/Book.pdf>
24. Fujita M., Krugman P. The New Economic Geography: Past, Present and the Future // Papers in Regional Science. — Wiley-Blackwell, 2004.

Приложение 1. Система уравнений 2

Северо-Западный ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_2

Инструменты: const CPI_2 EAP_2

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-17,1715	7,10933	-2,415	0,0157	**
CPI_2	0,206047	0,0635482	3,242	0,0012	***

Среднее зав. перемен	5,857143	Ст. откл. зав. перемен	1,543579
Сумма кв. остатков	16,51009	Ст. ошибка модели	1,172962
R-квадрат	0,466974	Испр. R-квадрат	0,422556

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_2

Инструменты: const CPI_2 EAP_2

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-1,29921e+07	4,74766e+07	-0,2737	0,7844	
EAP_2	296533	596137	0,4974	0,6189	
unempl_2	-827708	1,08604e+06	-0,7621	0,4460	

Среднее зав. перемен	2705398	Ст. откл. зав. перемен	1723390
Сумма кв. остатков	1,52e+13	Ст. ошибка модели	1176590
R-квадрат	0,656150	Испр. R-квадрат	0,593632

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_2

Инструменты: const CPI_2 EAP_2

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-1,63850e+06	1,69878e+06	-0,9645	0,3348	
GRP_2	0,319366	0,0463424	6,891	5,52e-012	***
CPI_2	13460,1	14216,9	0,9468	0,3438	

Среднее зав. перемен	729869,9	Ст. откл. зав. перемен	467234,9
Сумма кв. остатков	1,92e+11	Ст. ошибка модели	132234,1
R-квадрат	0,941355	Испр. R-квадрат	0,930693

Общая матрица ковариации (cross-equation VCV)

(корреляции выше главной диагонали)

1,1793	(0,684)	(0,026)
7,7523e+005	1,0877e+012	(-0,507)
3287,3	-6,1918e+010	1,3739e+010

Логарифм определителя = 49,8624

Приложение 2. Система уравнений 3

Южный ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_3

Инструменты: const CPI_3 EAP_3

	Коэффициент	Ст. ошибка	z	P-значение	
const	-33,3585	16,3088	-2,045	0,0408	**
CPI_3	0,380486	0,147182	2,585	0,0097	***
Среднее зав. перемен	8,785714	Ст. откл. зав. перемен	2,022511		
Сумма кв. остатков	34,15557	Ст. ошибка модели	1,687097		
R-квадрат	0,357702	Испр. R-квадрат	0,304177		

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_3

Инструменты: const CPI_3 EAP_3

	Коэффициент	Ст. ошибка	z	P-значение	
const	740351	2,70610e+06	0,2736	0,7844	
EAP_3	80355,2	29371,4	2,736	0,0062	***
unempl_3	-501218	152719	-3,282	0,0010	***
Среднее зав. перемен	1585132	Ст. откл. зав. перемен	1070718		
Сумма кв. остатков	4,36e+12	Ст. ошибка модели	629216,0		
R-квадрат	0,777947	Испр. R-квадрат	0,737574		

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_3

Инструменты: const CPI_3 EAP_3

	Коэффициент	Ст. ошибка	z	P-значение	
const	-2,09640e+06	1,03275e+06	-2,030	0,0424	**
GRP_3	0,451111	0,0306653	14,71	5,50e-049	***
CPI_3	18106,7	8918,52	2,030	0,0423	**
Среднее зав. перемен	624244,3	Ст. откл. зав. перемен	436289,8		
Сумма кв. остатков	1,79e+10	Ст. ошибка модели	40349,95		
R-квадрат	0,992847	Испр. R-квадрат	0,991547		

Общая матрица ковариации (cross-equation VCV)
(корреляции выше главной диагонали)

2,4397	(0,741)	(-0,237)
6,4521e+005	3,1107e+011	(-0,514)
-13228,	-1,0256e+010	1,2792e+009

Логарифм определителя = 47,158

Приложение 3. Система уравнений 4

Северо-Кавказский ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_4

Инструменты: const CPI_4 EAP_4

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-26,9429	21,3665	-1,261	0,2073	
CPI_4	0,399043	0,193249	2,065	0,0389	**
Среднее зав. перемен	17,15714	Ст. откл. зав. перемен	2,691215		
Сумма кв. остатков	69,47003	Ст. ошибка модели	2,406069		
R-квадрат	0,262168	Испр. R-квадрат	0,200682		

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_4

Инструменты: const CPI_4 EAP_4

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-625023	5,59272e+06	-0,1118	0,9110	
EAP_4	58508,3	67637,1	0,8650	0,3870	
unempl_4	-141367	86482,0	-1,635	0,1021	
Среднее зав. перемен	595838,2	Ст. откл. зав. перемен	417722,7		
Сумма кв. остатков	7,09e+11	Ст. ошибка модели	253803,5		
R-квадрат	0,767237	Испр. R-квадрат	0,724917		

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_4

Инструменты: const CPI_4 EAP_4

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-397690	132787	-2,995	0,0027	***
GRP_4	0,359302	0,0100315	35,82	5,90e-281	***
CPI_4	3389,36	1167,34	2,903	0,0037	***
Среднее зав. перемен	190968,3	Ст. откл. зав. перемен	144104,3		
Сумма кв. остатков	1,45e+09	Ст. ошибка модели	11491,90		
R-квадрат	0,994621	Испр. R-квадрат	0,993643		

Общая матрица ковариации (cross-equation VCV)

(корреляции выше главной диагонали)

4,9621	(0,747)	(-0,035)
3,7450e+005	5,0613e+010	(-0,015)
-794,88	-3,5002e+007	1,0376e+008

Логарифм определителя = 43,888

Приложение 4. Система уравнений 5

Приволжский ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_5

Инструменты: const CPI_5 EAP_5

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-16,1413	7,97878	-2,023	0,0431	**
CPI_5	0,210655	0,0720063	2,926	0,0034	***

Среднее зав. перемен	7,185714	Ст. откл. зав. перемен	1,342705
Сумма кв. остатков	13,68021	Ст. ошибка модели	1,067716
R-квадрат	0,416302	Испр. R-квадрат	0,367661

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_5

Инструменты: const CPI_5 EAP_5

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-6,75041e+07	1,80544e+07	-3,739	0,0002	***
EAP_5	1,17868e+06	238955	4,933	8,11e-07	***
unempl_5	-949138	352376	-2,694	0,0071	***

Среднее зав. перемен	4142207	Ст. откл. зав. перемен	2516987
Сумма кв. остатков	4,89e+12	Ст. ошибка модели	666749,6
R-квадрат	0,942930	Испр. R-квадрат	0,932553

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_5

Инструменты: const CPI_5 EAP_5

	Коэффициент	Ст. ошибка	z	Р-значение	
const	-1,49256e+06	633630	-2,356	0,0185	**
GRP_5	0,284225	0,00938037	30,30	1,15e-201	***
CPI_5	12052,0	5460,92	2,207	0,0273	**

Среднее зав. перемен	1019347	Ст. откл. зав. перемен	686154,9
Сумма кв. остатков	3,42e+10	Ст. ошибка модели	55771,63
R-квадрат	0,994419	Испр. R-квадрат	0,993404

Общая матрица ковариации (cross-equation VCV)
(корреляции выше главной диагонали)

0,97716	(0,519)	(-0,162)
3,0330e+005	3,4929e+011	(-0,099)
-7928,4	-2,8805e+009	2,4439e+009

Логарифм определителя = 47,8319

Приложение 5. Система уравнений 6

Уральский ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_6

Инструменты: const CPI_6 EAP_6

	Коэффициент	Ст. ошибка	z	P-значение	
const	-16,8621	7,07305	-2,384	0,0171	**
CPI_6	0,215216	0,0632293	3,404	0,0007	***
Среднее зав. перемен	7,192857	Ст. откл. зав. перемен		1,446196	
Сумма кв. остатков	13,83361	Ст. ошибка модели		1,073686	
R-квадрат	0,491211	Испр. R-квадрат		0,448812	

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_6

Инструменты: const CPI_6 EAP_6

	Коэффициент	Ст. ошибка	z	P-значение	
const	-4,75341e+07	2,11143e+07	-2,251	0,0244	**
EAP_6	839614	270306	3,106	0,0019	***
unempl_6	-777384	457034	-1,701	0,0890	*
Среднее зав. перемен	3830118	Ст. откл. зав. перемен		2230449	
Сумма кв. остатков	1,01e+13	Ст. ошибка модели		958861,0	
R-квадрат	0,853211	Испр. R-квадрат		0,826522	

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_6

Инструменты: const CPI_6 EAP_6

	Коэффициент	Ст. ошибка	z	P-значение	
const	-3,12400e+06	2,05833e+06	-1,518	0,1291	
GRP_6	0,341534	0,0380088	8,986	2,57e-019	***
CPI_6	25662,7	17213,7	1,491	0,1360	
Среднее зав. перемен	1052477	Ст. откл. зав. перемен		659606,8	
Сумма кв. остатков	1,52e+11	Ст. ошибка модели		117723,4	
R-квадрат	0,973190	Испр. R-квадрат		0,968315	

Общая матрица ковариации (cross-equation VCV)

(корреляции выше главной диагонали)

0,98812	(0,704)	(0,345)
5,9448e+005	7,2240e+011	(0,154)
35760,	1,3693e+010	1,0889e+010

Логарифм определителя = 49,5774

Приложение 6. Система уравнений 7

Сибирский ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_7

Инструменты: const CPI_7 EAP_7

	Коэффициент	Ст. ошибка	z	P-значение	
const	-23,5457	6,49925	-3,623	0,0003	***
CPI_7	0,294602	0,0581422	5,067	4,04e-07	***
Среднее зав. перемен	9,357143	Ст. откл. зав. перемен		1,717652	
Сумма кв. остатков	12,21680	Ст. ошибка модели		1,008993	
R-квадрат	0,681475	Испр. R-квадрат		0,654931	

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_7

Инструменты: const CPI_7 EAP_7

	Коэффициент	Ст. ошибка	z	P-значение	
const	-9,80487e+06	2,74077e+07	-0,3577	0,7205	
EAP_7	308062	374191	0,8233	0,4104	
unempl_7	-812110	347559	-2,337	0,0195	**
Среднее зав. перемен	2802798	Ст. откл. зав. перемен		1665567	
Сумма кв. остатков	8,79e+12	Ст. ошибка модели		894049,8	
R-квадрат	0,777311	Испр. R-квадрат		0,736822	

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_7

Инструменты: const CPI_7 EAP_7

	Коэффициент	Ст. ошибка	z	P-значение	
const	-2,55109e+06	1,55843e+06	-1,637	0,1016	
GRP_7	0,340392	0,0409559	8,311	9,47e-017	***
CPI_7	20176,1	12948,1	1,558	0,1192	
Среднее зав. перемен	656338,6	Ст. откл. зав. перемен		484007,0	
Сумма кв. остатков	2,61e+10	Ст. ошибка модели		48724,06	
R-квадрат	0,991426	Испр. R-квадрат		0,989867	

Общая матрица ковариации (cross-equation VCV)
(корреляции выше главной диагонали)

0,87263	(0,639)	(-0,088)
4,7340e+005	6,2804e+011	(-0,100)
-3558,2	-3,4355e+009	1,8653e+009

Логарифм определителя = 47,8394

Приложение 7. Система уравнений 8

Дальневосточный ФО

Уравнение 1: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: unempl_8

Инструменты: const CPI_8 EAP_8

	Коэффициент	Ст. ошибка	z	P-значение	
const	-24,2643	8,07753	-3,004	0,0027	***
CPI_8	0,292187	0,0723173	4,040	5,34e-05	***
Среднее зав. перемен	8,350000	Ст. откл. зав. перемен		1,625636	
Сумма кв. остатков	14,55496	Ст. ошибка модели		1,101324	
R-квадрат	0,576336	Испр. R-квадрат		0,541031	

Уравнение 2: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: GRP_8

Инструменты: const CPI_8 EAP_8

	Коэффициент	Ст. ошибка	z	P-значение	
const	-2,81580e+07	5,99256e+06	-4,699	2,62e-06	***
EAP_8	473823	83688,7	5,662	1,50e-08	***
unempl_8	-320851	82496,1	-3,889	0,0001	***
Среднее зав. перемен	1352369	Ст. откл. зав. перемен		892806,2	
Сумма кв. остатков	1,40e+12	Ст. ошибка модели		356403,2	
R-квадрат	0,874201	Испр. R-квадрат		0,851328	

Уравнение 3: 2МНК, использованы наблюдения 2000–2013 (T = 14)

Зависимая переменная: inv_8

Инструменты: const CPI_8 EAP_8

	Коэффициент	Ст. ошибка	z	P-значение	
const	-226135	1,76019e+06	-0,1285	0,8978	
GRP_8	0,390255	0,0729214	5,352	8,71e-08	***
CPI_8	1491,99	14941,7	0,09985	0,9205	
Среднее зав. перемен	468172,2	Ст. откл. зав. перемен		346551,5	
Сумма кв. остатков	7,53e+10	Ст. ошибка модели		82744,21	
R-квадрат	0,951940	Испр. R-квадрат		0,943202	

Общая матрица ковариации (cross-equation VCV)
(корреляции выше главной диагонали)

1,0396	(0,701)	(0,201)
2,2594e+005	9,9804e+010	(0,046)
15035,	1,0555e+009	5,3795e+009

Логарифм определителя = 47,0338

Приложение 8. Модель с фиксированными эффектами

Модель 7: фиксированные эффекты, использовано наблюдений - 120
 Включено 8 пространственных объектов
 Длина временного ряда = 15
 Зависимая переменная: l_GRP

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	4,16550	0,433704	9,604	3,42e-016	***
l_Invest	0,652606	0,0274004	23,82	5,22e-045	***
employed	0,000182936	5,48347e-05	3,336	0,0012	***
income	2,05368e-05	2,64505e-06	7,764	4,79e-012	***
Среднее зав. перемен					
	14,55991	Ст. откл. зав. перемен	1,093830		
Сумма кв. остатков		0,926895	Ст. ошибка модели	0,092215	
LSDV R-squared		0,993490	Within R-squared	0,986403	
LSDV F(10, 109)		1663,435	P-значение (F)	3,0e-114	
Лог. правдоподобие		121,5318	Крит. Акаике	-221,0635	
Крит. Шварца		-190,4011	Крит. Хеннана-Куинна	-208,6114	
Параметр rho		0,624317	Стат. Дарбина-Вотсона	0,558926	

Joint test on named regressors -

Тестовая статистика: $F(3, 109) = 2635,92$

p-значение = $P(F(3, 109) > 2635,92) = 1,55787e-101$

Тест на различие констант в группах -

Нулевая гипотеза: Группы имеют общие константы

Тестовая статистика: $F(7, 109) = 30,553$

p-значение = $P(F(7, 109) > 30,553) = 5,1772e-023$

Тест на нормальное распределение ошибок -

Нулевая гипотеза: ошибки распределены по нормальному закону

Тестовая статистика: $\chi^2(2) = 6,45194$

p-значение = 0,0397172

Приложение 9. Модель со случайными эффектами

Модель 8: Случайные эффекты (GLS), использовано наблюдений – 120
 Включено 8 пространственных объектов
 Длина временного ряда = 15
 Зависимая переменная: l_GRP

	Коэффициент	Ст. ошибка	t-статистика	P-значение	
const	4,78456	0,299784	15,96	4,81e-031	***
l_Invest	0,680674	0,0256414	26,55	4,15e-051	***
employed	6,62072e-05	9,98334e-06	6,632	1,10e-09	***
income	1,96017e-05	2,66508e-06	7,355	2,94e-011	***
Среднее зав. перемен	14,55991	Ст. откл. зав. перемен	1,093830		
Сумма кв. остатков	3,085563	Ст. ошибка модели	0,162396		
Лог. правдоподобие	49,37284	Крит. Акаике	-90,74568		
Крит. Шварца	-79,59571	Крит. Хеннана-Куинна	-86,21762		

Внутригрупповая дисперсия = 0,00850363

Межгрупповая дисперсия = 0,0189128

theta, использованная для квази-деусреднения (demeaning) = 0,826868

corr(y, yhat)^2 = 0,978507

Тест Бриша-Пэгана (Breusch-Pagan) -

Нулевая гипотеза: Дисперсия специфических для наблюдений ошибок = 0

Асимптотическая тестовая статистика: Хи-квадрат(1) = 252,601

p-значение = 7,03564e-057

Тест Хаусмана (Hausman) -

Нулевая гипотеза: ОМНК оценки состоятельны

Асимптотическая тестовая статистика: Хи-квадрат(3) = 8,53256

p-значение = 0,0361967

Приложение 10. Данные

Region	Year	GRP	Invest	employed	income
Центральный	2000	1841499	303920	17721,30	3231
Центральный	2001	2243525	349312	17747,00	4300
Центральный	2002	2878665	435812	17927,80	5436
Центральный	2003	3577143	563111	18056,80	7211
Центральный	2004	4617086	770408	18220,40	8991
Центральный	2005	6278359	964156	18357,40	11084
Центральный	2006	7965170	1225593	18464,90	13883
Центральный	2007	10208918	1779599	18732,10	17084
Центральный	2008	12674395	2278329	19016,90	19141
Центральный	2009	11405184	1928138	18601,40	22628
Центральный	2010	13444440	2099825	18619,20	24525
Центральный	2011	16062124	2458313	18710,30	27089
Центральный	2012	17432295	2961586	18814,10	30006
Центральный	2013	18975900	3287365	18894,70	33467
Центральный	2014	20820579	3435974	19008,30	34970
Северо-Западный	2000	578505	116663	6609,60	2269
Северо-Западный	2001	709025	168114	6628,60	3084
Северо-Западный	2002	886843	199102	6659,50	4010
Северо-Западный	2003	1091027	285159	6689,90	5403
Северо-Западный	2004	1474882	359562	6713,80	6938
Северо-Западный	2005	1799780	483266	6737,90	9045
Северо-Западный	2006	2198608	651271	6801,10	10971
Северо-Западный	2007	2770190	832478	6854,20	13282
Северо-Западный	2008	3388222	1040669	6835,50	15070
Северо-Западный	2009	3415871	933693	6764,90	17662
Северо-Западный	2010	3943054	1176969	6764,50	19741
Северо-Западный	2011	4785459	1374216	6760,20	21184
Северо-Западный	2012	5247509	1544437	6774,20	23422
Северо-Западный	2013	5586594	1258559	6768,00	26167
Северо-Западный	2014	5914797	1357860	6750,20	28572
Южный	2000	329695	108941	5764,70	1592
Южный	2001	426508	130499	5883,20	2165
Южный	2002	519019	142905	5966,30	2898
Южный	2003	616085	162522	5987,10	3747
Южный	2004	766851	202691	6031,40	4565
Южный	2005	936056	245104	6056,80	5783
Южный	2006	1195195	324543	6076,60	7292
Южный	2007	1577083	496911	6137,70	9350
Южный	2008	2001112	704183	6200,70	11525
Южный	2009	1994913	709382	6139,20	13056
Южный	2010	2337937	907961	6113,60	15031
Южный	2011	2777792	1079284	6130,90	16584
Южный	2012	3185420	1254959	6201,10	18864

Продолжение таблицы

Region	Year	GRP	Invest	employed	income
Южный	2013	3528190	1428562	6197,80	21842
Южный	2014	3920265	1277238	6161,20	24328
Северо-Кавказский	2000	105178	25964	2646,10	1130
Северо-Кавказский	2001	142442	37102	2688,50	1441
Северо-Кавказский	2002	174564	42817	2719,90	1922
Северо-Кавказский	2003	220170	49663	2750,40	2532
Северо-Кавказский	2004	275606	61647	2766,30	3466
Северо-Кавказский	2005	352070	93318	2865,60	4544
Северо-Кавказский	2006	457118	128950	2899,80	5972
Северо-Кавказский	2007	573220	199888	3180,00	7606
Северо-Кавказский	2008	728231	260241	3238,10	9780
Северо-Кавказский	2009	786671	267085	3251,90	11553
Северо-Кавказский	2010	891834	313412	3314,60	13249
Северо-Кавказский	2011	1066320	347504	3374,30	15050
Северо-Кавказский	2012	1209039	402809	3397,20	17167
Северо-Кавказский	2013	1359273	426575	3423,30	18900
Северо-Кавказский	2014	1587148	516921	3464,10	20692
Приволжский	2000	1036789	206781	14209,90	1726
Приволжский	2001	1292757	267845	14293,40	2319
Приволжский	2002	1483310	294507	14481,00	3035
Приволжский	2003	1807987	350622	14460,00	3917
Приволжский	2004	2284896	464094	14486,50	4787
Приволжский	2005	2799036	609499	14503,80	6220
Приволжский	2006	3513342	783640	14614,80	7996
Приволжский	2007	4330428	1148397	14687,90	9930
Приволжский	2008	5324051	1485341	14665,30	12351
Приволжский	2009	4922532	1279154	14402,50	13911
Приволжский	2010	5709470	1437472	14357,20	15697
Приволжский	2011	7050736	1702521	14295,00	17282
Приволжский	2012	7864342	2012877	14336,40	19663
Приволжский	2013	8571225	2228110	14217,00	21864
Приволжский	2014	9171075	2355973	14114,80	24020
Уральский	2000	866133	250731	5711,90	2744
Уральский	2001	1120820	330984	5836,40	3820
Уральский	2002	1335976	383378	5879,00	4791
Уральский	2003	1659322	445954	5993,30	6110
Уральский	2004	2234753	534467	6056,80	7413
Уральский	2005	3091363	593370	6093,10	9507
Уральский	2006	3720616	801479	6078,30	12038
Уральский	2007	4236325	1113151	6082,20	15025
Уральский	2008	4815668	1482552	6104,50	18685
Уральский	2009	4360451	1337857	6043,90	19769
Уральский	2010	5118918	1490849	6067,20	21586
Уральский	2011	6314341	1838272	6056,70	23908

Продолжение таблицы

Region	Year	GRP	Invest	employed	income
Уральский	2012	7098364	2037624	6062,30	26304
Уральский	2013	7648600	2094007	6053,40	28994
Уральский	2014	8001749	2322596	6037,10	30494
Сибирский	2000	687071	98647	8690,90	1933
Сибирский	2001	844142	135116	8702,80	2576
Сибирский	2002	991737	150109	8727,10	3373
Сибирский	2003	1209597	193614	8798,30	4351
Сибирский	2004	1631783	255399	8865,30	5267
Сибирский	2005	1951299	346106	8911,50	6680
Сибирский	2006	2442999	483721	8947,90	8346
Сибирский	2007	2990665	708951	9029,80	10286
Сибирский	2008	3442210	945556	9097,20	12873
Сибирский	2009	3391088	834593	8954,80	13525
Сибирский	2010	4131394	980474	9027,00	14892
Сибирский	2011	4802934	1219287	9009,70	16568
Сибирский	2012	5186809	1459474	9085,60	18474
Сибирский	2013	5535450	1377697	9061,00	20454
Сибирский	2014	6106913	1440980	9010,10	21490
Дальневосточный	2000	308802	53589	3162,20	2498
Дальневосточный	2001	391750	85743	3200,20	3304
Дальневосточный	2002	471106	113779	3212,90	4391
Дальневосточный	2003	561094	135723	3243,40	5788
Дальневосточный	2004	678448	216743	3266,50	7047
Дальневосточный	2005	826422	276291	3265,60	8892
Дальневосточный	2006	999073	330825	3290,70	11097
Дальневосточный	2007	1277127	436849	3315,30	13358
Дальневосточный	2008	1534868	584745	3315,40	15622
Дальневосточный	2009	1730519	686111	3304,40	18410
Дальневосточный	2010	2110721	787699	3313,30	20809
Дальневосточный	2011	2532572	1060505	3306,50	22870
Дальневосточный	2012	2702292	971353	3297,30	25504
Дальневосточный	2013	2808368	814456	3285,80	28929
Дальневосточный	2014	3222508	820142	3267,50	31974

Приложение 11. Данные

Центральный	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	1 841 498	2 243 525	2 878 664	3 577 142	4 617 086	6 278 359	7 965 169	10 208 917	12 674 395	11 405 184	13 444 440	16 062 123	17 432 294	18 975 900
Ур. ЭАН	65,9	65,1	65,7	65,6	66,3	66,6	66,9	67,9	67,9	68,1	68,3	69,2	70,0	69,9
Ур безработицы	7,8	6,0	5,1	5,1	4,7	4,3	4,0	3,1	3,6	5,8	4,6	4,1	3,1	3,3
Инвестиции в ОК	303 918	349 312	435 810	563 111	770 409	964 158	1 225 593	1 779 599	2 278 329	1 928 138	2 099 824	2 458 312	2 961 584	3 287 363
ИПЦ	120,7	120,7	115,3	107,2	110,5	110,5	109	112,2	113,3	109,4	108,9	106	106,9	106,7

Северо-Западный	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	578 504	709 025	886 843	1 091 027	1 474 882	1 799 780	2 198 608	2 770 190	3 388 222	3 415 871	3 943 053,7	4 785 458,7	5 247 509	5 586 593
Ур. ЭАН	66,4	66,2	66,6	67,0	67,4	68,8	70,0	70,7	71,1	71,2	70,8	71,4	71,5	70,9
Ур безработицы	9,6	7,6	6,2	7,0	6,0	5,4	4,9	4,1	5,0	6,9	5,9	5,1	4,0	4,3
Инвестиции в ОК	116 663	168 113	199 102	285 158	359 562	483 265	651 271	832 478	1 040 669	933 693	1 134 405	1 329 968	1 485 413	1 198 415
ИПЦ	121	121	118	110	112	111,2	108,6	112,6	114,1	108,5	109,2	105,8	106,1	106,6

Южный	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	329 695	426 508	519 019	616 085,3	766 851,3	936 055,9	1 195 195	1 577 083	2 001 112	1 994 913	2 337 937	2 777 792	3 185 419	3 528 190
Ур. ЭАН	62,8	61,6	62,5	61,0	63,3	62,5	63,7	64,8	65,1	65,1	64,5	65,1	65,6	86,8
Ур безработицы	12,9	11,5	9,6	11,4	9,6	8,4	8,2	7,0	6,4	8,6	7,6	7,0	6,2	8,6
Инвестиции в ОК	134 904	167 599	185 722	212 183	264 339	338 421	453 493	696 798	905 814	709 382	907 962	1 079 284	1 254 958	1 428 561
ИПЦ	117,7	112,1	111,3	112	112	113	109,5	112	113,6	109,2	109	106,1	106,6	106,6

Северо-Кавказский	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	105 178	142 442	174 564	220 170	275 606	352 070	457 118	573 220	728 231	786 670,9	891 834,3	1 066 319,6	1 209 039	1 359 273
Ур. ЭАН	61,2	59,1	59,6	58,6	59,5	62,2	61,6	63,1	63,5	64,5	63,9	64,6	65,3	65,8
Ур безработицы	20,4	18,7	17,4	17,2	18,8	17,1	22,6	19,2	15,7	16,0	16,5	14,5	13,1	13
Инвестиции в ОК	25 964	31 125	40 214	54 231	82 241	93 317	128 950	199 888	260 241,47	267 085,42	313 412,30	347 503,60	402 808,80	426 575
ИПЦ	116,1	114,2	110,1	108,3	110,5	111,2	109	113,3	116,1	109,9	110,6	105,2	106,6	106,1

Продолжение таблицы

Приволжский	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	1 036 789	1 292 757	1 483 310	1 807 987	2 284 896	2 799 036	3 513 342	4 330 428	5 324 051	4 922 532	5 709 470	7 050 735	7 864 342	8571225
Ур. ЭАН	65,9	64,8	65,4	64,5	65,2	65,8	66,2	66,8	66,9	67,6	67,8	68,4	68,6	68,1
Ур безработицы	9,8	8,4	7,7	7,7	7,9	7,4	6,5	6,1	6,2	8,6	7,6	6,5	5,3	4,9
Инвестиции в ОК	206 781	267 845	294 507	350 622	464 095	609 499	783 640	1 148 391	1 485 341	1 279 154	1 437 472	1 702 521	2 012 877	2 228 110
ИПЦ	121,2	115,3	109,6	110,5	112,3	110,2	108,7	113,1	113,3	107,9	109,3	106,2	106,4	106,3

Уральский	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	866 133	1 120 820	1 335 976	1 659 322	2 234 753	3 091 363	3 720 616	4 236 325	4 815 668	4 360 451,2	5 118 918,4	6 314 341,2	7 098 364,3	7648600
Ур. ЭАН	66,6	64,9	65,0	67,0	66,7	67,3	67,9	67,5	69,1	68,7	69,1	69,8	70,1	70
Ур безработицы	10,1	9,2	8,1	7,4	7,4	6,7	6,8	4,9	5,5	8,1	8,0	6,8	6,0	5,7
Инвестиции в ОК	250 731	330 984	383 378	445 954	534 467	593 370	801 479	1 113 151	1 482 552	1 337 857	1 490 849	1 838 272	2 037 624	2 094007
ИПЦ	122,2	118,5	116,4	114,2	110,5	111,7	110,2	110,9	112,6	108,8	109,8	106,4	106,4	106,2

Сибирский	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	687 071	844 142,2	991 737	12 095967	1 631 782	1 951 299	2 442 999	2 990 665,1	3 442 209,8	3 391 088,1	4 131 394,4	4 802 933,8	5 186 808,5	5535450
Ур. ЭАН	65,0	62,5	64,2	64,0	64,9	65,9	65,5	66,2	66,5	66,4	66,9	66,9	66,6	66,8
Ур безработицы	12,8	11,3	10,1	11,4	9,9	9,3	8,7	7,6	8,3	10,5	8,7	8,1	7,1	7,2
Инвестиции в ОК	98 647	135 116	150 109	193 614	255 399	346 105	483 721	708 950,60	945 556,16	834 593,33	980 472,00	1 219 287	1 459 474	1377696
ИПЦ	120,5	118,	117,5	116,3	112,8	110,5	108,6	110,8	112,9	108,4	107,9	106,3	106,7	106,1

Дальневосточный	2000г.	2001г.	2002г.	2003г.	2004г.	2005г.	2006г.	2007г.	2008г.	2009г.	2010г.	2011г.	2012г.	2013г.
ВРП	308 802	391 750	471 106	561 094	678 448	826 428	999 073	1 277 127	1 534 868	1 730 519	2 110 721	2 532 572	2 702 292	2808368
Ур. ЭАН	68,2	66,3	66,2	67,2	66,8	66,5	67,0	67,7	69,0	68,9	68,4	69,6	69,6	69,7
Ур безработицы	12,6	10,2	8,7	8,5	8,9	7,9	7,4	6,6	7,7	9,2	8,6	7,4	6,7	6,5
Инвестиции в ОК	53 589	85 743,40	113 779	135 7223	216 744	276 291	330 825	436 849	584 745	686 111	787 699	1 060 505	971 353	814 456
ИПЦ	118,4	117,2	115,6	114,3	115,2	113,3	108,8	109,6	113,6	109,7	107,7	106,8	105,9	106,6